DOCUMENT RESUME

ED 046 350 HE 001 940

AUTHOR	Eble, Kenneth E.
TITLE	The Recognition and Evaluation of Teaching.
INSTITUTION	Project to Improve College Teaching, Salt Lake City, Utah.
SPONS AGENCY	Carnegie Corp. of New York, N.Y.
PUB DATE	Nov 70
NOTE	117p.
AVAILABLE FROM	American Association of University Professors, One Dupont Circle, N.W., Washington D.C. ($1.00)
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	College Teachers, *Effective Teaching, *Evaluation, Evaluation Methods, Faculty Promotion, *Higher Education, Student Participation, *Teacher Evaluation, *Teaching Quality

ABSTRACT

This monograph brings together information and ideas about the recognition and evaluation of effective teaching in the college and university. The materials were gathered in 4 ways: (1) a search of the literature; (2) solicitation of information about systematic attempts to recognize and evaluate teaching: (3) visits to individual campuses and discussion of evaluation as part of these visits; and (4) a conference on April 10, 1970 in Washington, D.C. in which major issues and problems concerning this topic were discussed by students, faculty and administrators from a wide range of institutions. Chapter 1 discusses the need to improve college teaching, some of the reasons why teaching is neglected in favor of research, shaping of professional values, and the place of evaluation. Chapter 2 deals with ways of evaluating teacher effectiveness. Instruments and procedures for student evaluation of effective teaching are discussed in Chapter 3. Chapter 4 deals with the impact of student evaluation on the improvement of teaching, and Chapter 5 discusses the role of student evaluation in faculty review and advancement. Documents relating relevant experiences at individual universities, student evaluation forms, and a bibliography conclude the report. (AF)

ED0 46350

The Recognition and Evaluation of Teaching

by

Kenneth E. Eble, Director
Project to Improve College Teaching
and
The Committee on Evaluation

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Project to Improve College Teaching
1259 East South Temple, Salt Lake City, Utah 84110

HE 001 940

Price $1.00

The Project to Improve College Teaching

is jointly sponsored by the American
Association of University Professors and
the Association of American Colleges
and supported by a grant from
the Carnegie Corporation.

Advisory Board

Bruce Dearing, President, SUNY at Binghamton
Edward D. Eddy, President, Chatham College
Winston W. Ehrmann, Dean, Cornell College
John Gillis, Executive Associate, Association of American Colleges
Harry D. Gideonse, Chancellor, New School for Social Research
Wilbert McKeachie, Chairman, Psychology, University of Michigan
Neill Megaw, Chairman, English, University of Texas
Rosemary Pierrel, Dean, Pembroke College
Jim Sutton, Student, University of Iowa
Robert Van Waes, Associate Secretary, American Association of University Professors
Henry T. Yost, Professor of Biology, Amherst College

Additional copies of THE RECOGNITION AND EVALUATION OF TEACHING are available from The Project to Improve College Teaching (until September 1971) and from the American Association of University Professors, One Dupont Circle, Washington, D. C. 20036, at $1.00 per copy.

First printing: November, 1970

CONTENTS

PREFACE vii

1. Recognition of Teaching 1

2. Evaluation and the Improvement of Teaching 8

3. Student Evaluation Instruments and Procedures 21

4. Impact of Student Evaluation 34

5. Student Evaluation and Faculty Review 42

APPENDIX

1. Report of the President's Committee for a Pilot Study in Student Evaluation at Princeton 50

2. *From* Course Evaluation Booklet, Princeton University 59

3. Student Assessment of Teaching Effectiveness, by Thomas Langen, University of Washington 62

4. University of Washington Survey of Student Opinion of Teaching 68

5. University of Washington Report of the *ad hoc* Committee on Student Evaluation of Teaching 70

6. How Visitation Came to Carnegie-Mellon University, by Arthur Eastman 75

7. *from* Effective University Teaching and Its Evaluation, University of California at Davis 90

BIBLIOGRAPHY 101

Participants in Conference on Evaluation
sponsored by
The Project to Improve College Teaching, Washington, D.C., April 10, 1970

Walter Benjamin, Professor of Religion, Hamline University
Frank Chiarenza, Dean, School of Arts and Sciences, University of Hartford
Beekman Cottrell, Director of Undergraduate Studies in English, Carnegie-Mellon University
William Danner, Student, Harvard Graduate School of Business Administration
Frank Dilley, Associate Provost for Instruction, University of Delaware
Linda Douglass, Student, University of Tennessee
Arthur M. Eastman, Chairman, Department of English, Carnegie-Mellon University
Duane Ellison, Assistant Professor of History, Montgomery College
Kelby Fletcher, Student, The American University
Jerome J. Hanus, Associate Dean, School of Government and Public Administration, The American University
James Harris, Student, Princeton University
Milton Hildebrand, Professor of Zoology, University of California at Davis
Robert Hollander, Associate Professor of Romance Languages, Princeton University
Herbert Horowitz, Assistant Professor of Psychology, Wake Forest University
Kenneth L. Janzen, Acting Dean, Hamline University
Gordon Jensen, Chairman, Department of History, University of Hartford
Gwendolyn Jensen, Assistant Professor of History, New Haven College
H. Paul Kelly, Director of Testing, University of Texas at Austin
Richard L. Larson, Professor of English, University of Hawaii
Jonathan Lewis, Student, University of California at Davis
Neill Megaw, Chairman, Department of English, University of Texas at Austin
Thomas Mullen, Dean, Wake Forest University
Ohmer Milton, Director of Learning Center, University of Tennessee
J. Alden Nichols, Professor of History, University of Illinois at Urbana
William Rienstra, Instructor in Physics, University of Texas at Austin
Herman Sirois, Student, University of Hartford
Robert Van Waes, Associate Secretary, AAUP
Ronald H. Wenger, Assistant to Provost for Academic Planning, University of Delaware
Gerald Whitlock, Professor of Management and Organizational Psychology, University of Tennessee
Robert C. Wilson, Research Psychologist, Center for Research and Development in Higher Education, University of California at Berkeley
John Wriston, Professor of Chemistry, University of Delaware

Preface

This monograph brings together information and ideas about the recognition and evaluation of effective teaching in the college and university. This is one of the three major concerns of the Project to Improve College Teaching, created by a joint committee of the American Association of University Professors and the Association of American Colleges and supported by a grant from the Carnegie Corporation. The project began its two-year program of action and study in September, 1969. The materials in this monograph were gathered in four ways: (1) a search of the literature; (2) solicitation of information about systematic attempts to recognize and evaluate effective teaching; (3) visits to individual campuses and discussions of evaluation as part of these visits; and (4) a conference on April 10, 1970 in Washington, D.C., in which representatives from institutions identified with this aspect of the project's work discussed major problems and issues.

The April conference included students, faculty, and administrators from a wide range of colleges and universities. The general agreements reached covered most of the questions about the advisability, make-up, and uses of student evaluations. The group agreed that "wisely-formulated, wisely-administered, and wisely-used, student evaluations are useful to improving teaching," and also agreed that such input should be incorporated into the means of reaching decisions on retention, promotion, salary, and tenure. The chapters following do not attempt to record the conference discussions, but their substance owes much to that conference and to discussions at campuses from which many of the participants came. A draft of the manuscript was circulated to members of the conference to gain the benefits of their suggestions and revisions.

The material in the Appendix makes available some of the very useful documents received by the project. They should offer guidance in the particulars of instituting and administering evaluation procedures and in comprehensive plans for recognizing and evaluating effective teaching. By now, student evaluation of courses and teachers rests upon an impressive body of research data and upon their use in hundreds of colleges and universities. The answer to whether such evaluations result in improvement of teaching can only come from careful examination of their effects on individual campuses. It is hoped that this monograph will encourage such appraisals.

KENNETH E. EBLE

1. Recognition of Teaching

Just how good college teaching is throughout the country and how good it should be are two questions which have no exact answers. The complaints about the quality of undergraduate instruction are both current and chronic. Few recent speeches have made as much stir as William Arrowsmith's address to the 1966 annual meeting of the American Council on Education. "At present," he charged there, "the universities are as uncongenial to teaching as the Mojave Desert to a clutch of Druid priests. If you want to restore a Druid priesthood, you cannot do it by offering prizes for Druid-of-the-year. If you want Druids, you must grow forests."

In 1964 to 1966, W. Max Wise visited some thirty-five graduate schools and had the opportunity to examine informally the training of graduate assistants as future teachers. Though he found some important exceptions, his general conclusion was that "the universities, in their need to staff elementary courses, have thoroughly abused and prostituted the chief means of training prospective college teachers — the teaching assistantship."

A rare glimpse of teaching as it goes on inside the classroom was afforded by Paul Klapper, president emeritus of Queens College in 1949. In examining undergraduate instruction as it faced great growth in the size of individual classes and institutions, President Klapper spent two to three weeks each at six large universities. Three-fourths of his time was spent in visiting classes. His sketches (*Journal of General Education,* 3:228-44, 1949) do not give a very complimentary picture of undergraduate teaching. "Aimlessness," he concluded "is the most important single cause of ineffective teaching." His recommendations are as pertinent now as they were then: (1) the need for better preparation of graduate students as teachers, and

(2) effective in-service programs for improving the quality of instruction. "Generally speaking," he wrote, "the actual achievement of a college, an undergraduate institution, does not rise appreciably above the level of classroom performance."

How good is college teaching?

During the course of my visits to thirty campuses this past year, I have raised the general question of how good college teaching is and how good it should be with many faculty members and students. To give some definition to these informal inquiries, I have asked, at least half-seriously, how teaching on an individual campus might compare with baseball averages and whether reaching such a level of performance might be a reasonable expectation. Should a college staff maintain a teaching average about as high as a good major league team's won-loss record? Would an individual teacher be a power teacher through the season if he maintained a .300 average in the classroom? Or, calculating the number of outstanding teachers against the total number encountered, would a student feel he was getting fantastic instruction if he broke .400?

I think I would feel better about college teaching if I thought it approached baseball averages. Competent teachers are undoubtedly in fair supply on college faculties, but these have their baseball counterpart in the journeymen performers who provide the stuff for winter trades. Outstanding teachers in the colleges are probably less generously distributed than outstanding baseball players in the major and minor leagues. Moreover, it is the sterling performers, in baseball as in teaching, who have a disproportionate effect upon the young.

Such generalizations, I admit, are risky. No one knows how good or how bad college teaching is. This project proceeds from the general premise that almost everything needs improving and for the specific reason that those responsible for framing the project thought there was a need for improvement of college instruction. My own subjective impressions make me think that undergraduate teaching today is in need of forceful and continuing attention. But these feelings about the need for improving instruction cannot claim to be based on exact evidence. Rather, they probably arise from our perceptions of the nature of college teaching and an awareness of particular conditions which affect undergraduate insruction.

Perceptions of the nature of teaching

As to the former, teaching may suffer most from being the one thing almost everyone in the academic profession is expected to do. Teaching at its best is a great art, and great art of any kind is rare. Engaged in as a mass enterprise, for faculty as for students, teaching has difficulty rising above a kind of middle level that hits the students somewhat above where they sit down and well below where they think.

In the second place, teaching, for all its endless verbalizing, is a silent, secret art. Today's doctor of philosophy has a very private practice, enclosed by the specialized courses he can call his own and protected by academic traditions and superstitions. Among these are beliefs that the Ph.D. is a license to teach; that scholarly assiduity ensures good teaching or makes up for bad; that the popular teacher can't be profound and the profound one popular; that teaching can't be taught; and that, however deficient a professor may appear, he will turn out to be, for some student, some time, a superior teacher.

In the face of such beliefs, it is no wonder that teaching goes unrecognized. The effects of teaching are personal, slow to work, and slow to be discovered. Henry Adams' opinion that "the teacher affects eternity" has an impressive ring, but how does eternity record itself at deliberations of a promotions committee? It is not unimportant that teachers live by the long-range and undisclosed impact they make, but that does not set aside the fact that teachers need more immediate rewards, more tangible evidences of the worth of what they are doing.

It is equally important that students perceive teachers as other than mere parts of an institutional mechanism. The plea for "community" in higher education is not mere cant. The relationship between teacher and student is vital to teaching, and the genuine concern an institution shows for teachers and teaching is a direct measure of its concern for students and learning. Recognition of good teaching is not only a matter of giving continuing attention to it, but of trying to increase the worth of doing it well. Although there are many means of recognizing and valuing teaching within both the individual's and institution's grasp, ultimately we are facing the values abstractly held by the profession and given concrete form in institutional and individual practices.

Teaching, research, and service

A faculty member's responsibilities are thought of principally in terms of teaching, research, and service. Professional recognition as well as movement up the institutional ladder comes through achievements in one or all of these areas. Higher education is trinitarian in word if not in deed. Few universities conceive of themselves as teaching institutions or research institutions or even primarily dedicated to community service. Instead of a dozen institutions genuinely devoted to and actually fulfilling the aim of expanding the frontiers of knowledge, we have hundreds with such an objective as part of total aims. Conversely, almost all collegiate institutions have as a major function quantities of teaching which is far removed from the frontiers of knowledge. The generally accepted ways of defining institutional goals shape hiring, retention, and promotion policies within the departments and colleges.

A disproportionate attention to research is most often cited as the specific cause for the neglect of teaching. Service is what a faculty member may gripe about, dodge around, or even devote himself to; it takes his time and may get in the way of class preparation, but it does not pose as a superior and compelling rival to teaching. The arguments about teaching versus research have no easy resolution. In some disciplines and at some levels of study, research and teaching go hand in hand. Elsewhere, research may or may not be compatible with the individual professor's function as a teacher. Martin Trow flatly states that at the large state universities "the majority of university teachers are certainly not interested primarily in teaching." My own review of much writing on this subject leads me to believe that both research and teaching are too narrowly conceived. Research might better be called "search," a term more accurately descriptive of an activity all college and university professors must engage in. Research is a narrower term, describes a more specialized activity, and does not accurately define the wide range of creation and exploration, analysis and synthesis, speculation and reflection, appropriate to students and faculty in the college and university. Teaching, as it describes the principal activity of most professors, deserves to be regarded as something more than mere transfer of information or inculcation of skills. Teaching cannot maintain equal status with research until it ceases being regarded as an activity requiring less skill, less study, and less expenditure of intelligence and energy.

Recognition of teaching in this respect is part of the general problem, for individual institutions as for faculty, of upholding a unified ideal of teaching, research, and service and yet being able to give one or another first claim to attention. In the majority of undergraduate institutions, for example, surely teaching should have the highest priority. Were such priority given, much of the vague malaise about the teacher's low estate and the evils of publish or perish would diminish if not disappear. If priorities cannot be given, then strong reinforcements of more than a single value might be equally effective. As Paul Saltman, the provost of Revelle College, University of California at San Diego, recently said: "We have two mottoes on our campus: "Publish or perish: Teach or travel."

Shaping of professional values

The values of the individual professor are being shaped all the way through his formal education. Allan Cartter locates the root of current attitudes toward research and teaching in the graduate schools:

> The young faculty recruit is an impressionable person, and the two most important periods in his life are those that immediately precede and follow his getting the doctorate. Typically, as a graduate student, he pursued a degree in a large university which was itself caught up in catering to many markets and serving many masters, and therefore gave little attention to any one. It is puzzling that all two-hundred-odd graduate schools in this country assume this stance, and unfortunate that no single one has ever devoted itself exclusively to seeking students intent on an academic career and doing a really first-rate job of training them as teachers.

Too little is done to give recognition to teaching at all stages in the faculty member's career: in the graduate program, in the treatment of graduate assistants, in the ways of inducting new teachers into full-time positions, and in supporting and sustaining teaching of the fully established faculty member. The attitudes we have been talking about exert a powerful impact. The realities of campus life are just as powerful: graduate assistants are poorly paid and inadequately trained; grants, awards, released time, leaves in support of teaching are not widely available; extra outside or institutional funds are not often forthcoming in support of teaching.

Graduate education, in growth, money expended, prestige, and perhaps even in accomplishments, has dominated higher education since World War II. For most of the past decade, young faculty members coming out of the graduate schools have had their choice of positions. Institutions, hard-pressed to meet undergraduate enrollments, have characteristically struck bargains between staffing necessary courses and furthering the specialized scholarly pursuits of faculty members. As a consequence, graduate work has flourished in almost all large institutions, and teaching undergraduate classes has been something of the price faculty members as well as institutions have had to pay. The growth of universities in the number of students and in their capacities as research and service institutions has created conditions unfavorable to undergraduate teaching in many ways. Changes in both climate and conditions will have to come about before teaching gains a proper place.

But teachers teach, students learn, in bad weather as well as good. Some measures can be taken within the given conditions that might increase recognition of teaching, both in its visibility and in rewards. The intention of this monograph is to provide materials useful to refining the procedures by which teaching is evaluated in itself and as part of the professor's total competence. Teaching comes into prominence to some degree as departments colleges, and universities examine their criteria for good teaching, their procedures for reviewing it, and their stated expectations with regard to faculty performance. Such examination is necessary if reward is to follow recognition and if a college community as a whole is to perceive the high place of teaching in the reward structure.

The place of evaluation

The recognition and evaluation of effective teaching as discussed here have different applications for different colleges and universities. Some may shy away from formalizing the means by which a faculty member's teaching competence is judged. Some may be reluctant to give students a voice in the decisions which affect a faculty member's career. Others may accept the idea of gathering students' opinions about courses and teaching, and find that computer-processed questionnaires yield useful results. Any general practices need to be adapted to an individual institution's concept of itself. But this does not set aside the need in all colleges and universities to provide ways that the students can be heard with respect to teaching.

Teaching and student interaction

Why is this so? The most important reason is that teaching which disregards the interaction between student and teacher hardly justifies the holding of formal classes. We possess everything necessary to replace the information-dispensing teacher at a stroke. Perhaps the classroom and tutoring and counseling aspects of teaching will all diminish as educational technology increases the opportunities for students to learn on their own. Students who have experienced the kind of teaching which develops the individual's capacity and inclination to learn on his own are justifiably impatient with much of conventional college teaching. If teachers are to claim a worth at all, it must be in a large sense because of the interactions between teacher and student which move the student to become a self-motivating learner.

This basic necessity is as important on one campus as on another. Ideally, the teacher and students establish the necessary relationships by virtue of their coming together in class and out. But in the real world of higher education, to a great degree, such relationships depend upon the willingness of a competent scholar to expend energy and imagination and intellect on teaching itself. The conditions favorable to such teachers are not strongly present within the profession, nor are they widely established in individual colleges or universities. That is why specific means of giving conscious attention to teaching are so necessary.

It may not be in the various forms of evaluation discussed here, but it must be in some form that (1) attempts to reach all teachers and all students; (2) is continuing in its effects rather than sporadic; (3) focuses upon the interactions between the individual teacher and the individual student; and (4) gives the student responsibility for and active participation in his own education.

Other measures in support of teaching

Departments, colleges, and the university itself can take other measures which would increase the recognition given teaching. Grants for purposes directly affecting teaching can be provided as well as research funds. Deans and department chairmen can be active in encouraging a wide range of individual activities in support of effective teaching and in providing for campus- or department-wide consultation, discussion, and conferring to give teaching wider visibility. In a large university, an office for teaching and learning can be as important as an office in support of research. Relevant research in teaching and learning can be a part of such an office, not so much in fostering it but in providing ways by which research can make its impact on actual practices. Many campuses already support annual awards for superior teaching. Such awards in themselves are forms of recognition, even though they arouse legitimate questions as to whether funds and energy spent in this way might be better spent upon other activities.

What is most important to any means of recognizing teaching is the promise it gives of steady and continuing attention. Not all members of a faculty will flower as teachers in the most beneficent climate. Nor are all effective teachers dependent upon outside support to maintain their effectiveness. But whatever else an institution does to foster good teaching it should provide ordinary and accessible means by which a teacher can develop, and ordinary and effectively-operating means by which he can prosper.

2. Evaluation and the Improvement of Teaching

Evaluation is a loaded word. Faculty members are no different from other human beings if they stiffen slightly at the prospect of being evaluated. How does one go about evaluating as complex and personal an act as teaching, anyway? If we could evaluate teaching, would it lead to improvement? And how can we either evaluate or improve teaching when we don't know what good teaching is?

What we know about teaching effectiveness

These are troubling questions for the college professor, the moreso because faculties have had so few reservations about evaluating the learning of students. The questions are not, however, quite without answers. We know, for example, a good deal about effective and ineffective teaching. Certainly the hundreds of thousands of teachers who appear before students every day have some sense of the effectiveness or ineffectiveness of what they are doing. The individual teacher knows, in the pragmatic sense, that this direction is better than another, this approach has worked in the past and may work today, that these acts and attitudes on his part seem to invite the student in and incline him, at the least, toward learning.

It is Socratic wisdom that the mark of the knowing teacher is that he *knows* very little. Of the teaching process itself, he may only *know* that he must be constantly ready to drop old strategies and adopt new ones. But this *knowing* is part of the essential wisdom of any man or woman on the way to becoming educated: that few things are certain, that time and events require a continuing effort to recast what is known and to seek new ways of knowing as well as new knowledge itself.

What most college teachers are not aware of is the great number of careful scholarly investigations which have tried to find out what does constitute effective and ineffective teaching. The extensive bibliography accompanying this monograph surveys only one area of study — evaluation of the teaching process itself — and for the most part includes only those published materials based upon experience with or investigations of student observations of college and university teaching. Were the results of these investigations at wide variance with one another, the skeptic might find support for his distrust of knowing anything about the subject. But the investigations are not at variance, and whether one is looking for the general qualities of the effective teacher or for more specific characteristics of effective teaching, research studies provide reasonably consistent answers. Such findings tend also to square with practical experience.

The general qualities which distinguish the superior teacher have been set down by almost every teacher with a keen interst in teaching. Julius Taylor, head of the Physics department at Morgan State, puts it in the form of an equation: $C_1 + C_2 + C_3 + E$ = Effective Teaching, where C_1 = Competence, C_2 = Concern or Compassion, C_3 = Commitment, and E = Enthusiasm. In response to a request from one of the colleges I visited this year, I came up with *competence, caring, energy, imagination,* and *a sense of proportion* as prime qualities. Many years ago, a University of Chicago study made by students and faculty arrived at *knowledge and organization of subject matter, skill in instruction,* and *personal qualities* as the classes of attributes most important to the good teacher. For all that the diversity of students and teachers draws attention to the variables that make up good teaching, such variables are best perceived as diverse personal manifestations of commonly agreed-upon qualities.

If we turn to what teachers do in class, we are again confronted not with a hopelessly large and varied collection of attributes, but with a relatively small number of characteristics which relate to a teacher's effectiveness. The extensive research on student questionnaires for rating classroom performance has a high degree of uniformity. A factor analysis of forty-one items most frequently mentioned in the literature is the basis for the ten items which appear on the current University of Washington questionnaire: 1. *Interprets abstract ideas and theories clearly;* 2. *Gets me interested in his subject;* 3. *Has increased my skills in thinking;* 4. *Has helped broaden my interests;* 5. *Stresses important material;* 6. *Makes good use of examples and illustrations;* 7. *Has motivated me to do my best work;* 8. *Inspires class confidence in his knowledge of subject;* 9. *Has given me new viewpoints or* *appreciations;* 10. *Is clear and understandable in his explanations.* A recent study at the University of California at Davis groups the components of effective teaching as perceived by students into five scales: 1. *Analytic/synthetic approach;* 2. *organization/clarity;* 3. *instructor-group interaction;* 4. *instructor-individual student interaction;* and 5. *dynamism/enthusiasm.* The most significant items in each scale are: 1. *Discusses points of view other than his own;* 2. *Explains clearly;* 3. *Encourages class discussion;* 4. *Has a genuine interest in students;* 5. *Is a dynamic and energetic person.* Wilbert McKeachie at the University of Michigan reports a study which shows that psychology instructors whose students do best on tests of psychological thinking tend to be described as *listening attentively to what class members say, friendly, permissive and flexible, clear in explanations, and skillful in observing student reactions.*

This kind of information (and there is a great deal available) is at the first level of perception about what constitutes good teaching, but it is a necessary perception if one is to consider fairly evaluation as a means of improving teaching. The profession does have substantial information about the effectiveness of both teachers and teaching. There is no great disparity between the findings of one research study and another, between what research and informed opinion have to say about effective teaching, and (at the point of teaching performance) between what students and faculty say. What is wanting, in improving teaching as in almost any endeavor, is a fruitful correspondence between research and practice, knowing and doing. Human beings do not easily perform at the level of which they are capable nor in the light of what is known or could be found out.

Ways of judging teaching effectiveness

If one grants that we do know something about the effectiveness of teaching, are there ways of making a useful evaluation of its effectiveness? College teaching has not escaped evaluation of both a formal and informal kind. Academic rank and tenure are tied to somewhat common practices throughout the country's colleges and universities. The 1966 Astin and Lee study of the evaluation of teaching effectiveness made clear that the judgment of a chairman or dean, supported or confirmed by the opinions of departmental colleagues, is the most commonly-used means of evaluating an individual faculty member's teaching competence. Firsthand knowledge of a faculty member's teaching plays a small part in this kind of evaluation, and scholarly research and publication and other inferential evidences of teaching competence are given substantial weight. Astin's and Lee's comment seems to be a fair criticism of current practices: "If the ultimate measure of the teacher's effectiveness is his impact on the student — a view which few educators would dispute — it is unfortunate that those sources of information most likely to yield information about this influence are least likely to be used."

Classroom visitation

What are some of the ways by which more direct and significant information about teaching competence might be obtained? Classroom visitation is an obvious means, but at no time has it been a very widespread practice. My guess is that it is somewhat less used now than it has been in the past. The growth in institutions, for example, has probably made it more difficult to find the time for systematic visitation, and the current climate within higher education may be unfavorable even to the helpful, parental kind of visiting. The Astin and Lee survey gives some statistical support to these conjectures. But whether visitation is on the decline or not, it is a means of evaluation which gains its strong proponents as well as arouses widespread faculty resistance. Though individual faculty members in many colleges may occasionally visit another professor's classes, few colleges have formal boards of visitors or a tradition which makes visiting classes a natural, casual, and revealing activity. Arthur Eastman's account of his attempt to introduce classroom visitation at Carnegie-Mellon (see Appendix) raises many of the important questions about the possible use and effectiveness of classroom visits in evaluating teaching.

Self-appraisal

Self-scrutiny is another means of gaining information useful to evaluating teaching effectiveness. We have probably passed the peak of popularity for institutional self-studies, and with it, perhaps, the peak of interest in faculty self-evaluation as a means of improving instruction. The self-evaluation which precisely sets forth a teacher's objectives and which with equal precision describes and analyzes how these objectives might be achieved can be a brilliant contribution to our knowledge about teaching. Courses and students can benefit from such exercises fully as much as the individual teacher. Some institutions ask for self-appraisals as part of the basic data for judging competence, for the confidential departmental or college review of individual faculty members is a source of some unease and not always a very good means of assisting in a faculty member's development. Some departments appoint a committee to examine a faculty member's qualifications and discuss with him his perception of what he is trying to do and how well he is doing it. At Hamline University, the faculty member may submit a formal self-appraisal in addition to either student evaluations or the reports of divisional visitors as supporting material for promotion or salary increases. Obviously, there are many ways in which self-appraisal can be profitably used.

Course materials and data

A third means by which institutions might judge teaching effectiveness relies upon the materials used in the conduct of formal instruction and the various data which measure student performance in a faculty member's classes. Course syllabi and examinations are readily available and probably get looked at by department chairmen either in passing or at the time of formal review. Student credit hours handled, kinds of courses taught, enrollments in elective classes, grade distributions, student performance on examinations, supervision of independent study or projects, are measures which are readily accessible to department chairmen or committees. It is probably safe to say that much caution is exercised in drawing upon this kind of evidence. Gaining consensus about the use of such materials, trying to gather comparable data, and attempting to set norms or make comparisons pose great difficulties.

Student opinion

Informal student opinion, in the Astin and Lee survey, ranked fifth (just below scholarly research and publications) as a source of information for evaluating teaching effectiveness. Systematic student ratings were well down the list both in the 1961 survey and the 1966 one. These responses do not reflect the *presence* of systematic ratings on campuses, but the actual *use* made of them by departments. The number of student rating systems in use is difficult to determine, for the numbers have grown rapidly in the last decade and no accurate survey has been made to determine just how many are currently in use. Certainly over half of all institutions and a much higher percentage of major institutions have or have had some form of student evaluation. Assessment of student opinion of teaching effectiveness has been conducted continuously at the University of Washington since 1925. H. H. Remmers' studies of teaching evaluations began appearing in the late nineteen-twenties. Harvard's "Confidential Guide" was initiated in 1924. Though the student opinion questionnaire is probably the most common means of getting student input, other means have been used by some researchers or in specific colleges or universities. The exit interview in which all or some members of a graduating class are invited to review their work in a college or department is one technique. Another is a sampling of alumni, either by questionnaire or by an invitation to make an essay response. At Wayne State, a small number of students who are "experientially in touch" sit in on classes for a week and write course evaluations. Selected students or groups of students have been used elsewhere to provide information about courses and teachers in both formal and informal ways. Flanagan's "critical incident" technique is one of a number of attempts to provide better criteria for the evaluating of teaching performance.

Recording devices

Finally, the availability of recording devices has opened up many possibilities for observation and analysis which might be used in judging teaching competence. At Harvard, Kiyo Morimoto has assembled a remarkable collection of audio tapes of lectures and tutorials by teachers at Harvard and elsewhere in many different subjects. These have been used to acquaint fellows with different teaching styles and to sharpen their awareness of all that contributes to effective or ineffective teaching. The availability of television tape gives greater range to this kind of observation and analysis. Such devices afford a maximum opportunity to reduce the threatening or coercive aspects of observation and evaluation. Their use as a standard practice in the preparation of beginning teachers could arouse an interest in teaching itself which might be sustained and communicated to others throughout a faculty member's career.

Team teaching

If we add team teaching to the use of these recording devices, we extend the possibilities of direct observation by colleagues without intrusion into the classroom. Team teaching, ordinarily thought of as benefiting the student, may be of equal benefit to the teacher. The chance to observe other teachers in action and to benefit from being observed is almost ideally present in team teaching. In addition, team teaching can be a way of engaging a department's interest in teaching and in affording all members of the department a wider acquaintance with different teaching styles.

The importance of student opinion

There are many sources of information about teaching effectiveness and many ways of bringing that information to bear upon the evaluation of teaching. For these reasons, alone, the decisions reached as to a faculty member's teaching competence are likely to be based upon a number of independent judgments drawing upon various kinds and sources of information. This is probably as it should be. What seems to be most lacking in current practices is carefully accumulated information about a teacher's actual performance. Student opinion is of particular importance here because it represents an important addition to the data customarily used to judge faculty competence. It is the one source of direct and extensive observations of the way teachers carry out their daily and long-range tasks.

But before going on to discuss specific aspects of student evaluation, let me acknowledge the gap that exists between observations of teaching performance (whether by students, administrators, or faculty) and what the students actually learn. Measuring a student's learning would appear to be the one sure way of quantitatively assessing teaching effectiveness. What appears simple, however, is not really so. In practice, college and university instruction proceeds within a rigid framework of performance evaluation, though admittedly of the student's performance rather than of the faculty's. Its imperfections have been the source of long and increasing complaints about grading, the course system, the assigning of credit hours, and the cutting up of knowledge to fit the system. The last decade has been a remarkable one for the general lifting of these restrictions upon both teaching and learning. Measuring teaching effectiveness by the "before and after" achievements of students would necessitate relying upon measurements of students' abilities and achievements which are themselves under question.

Beyond that lie the complexities of determining just what part the individual teacher plays in a student's learning. In some ways, current research into college students' growth and development tends to diminish the importance of formal class work. Careful investigation seems to support what common sense and humility already tell the teacher: that his efforts play a small part in what the student learns and what he does with that learning. But the implication of such findings is not that we should give up teaching (and few teachers are going to do so regardless of the weight of evidence) but that we should look at teaching with full awareness of its broader contexts. The validity of the relationship between teacher and learner has the evidence of individual history as well as the history of civilization behind it. We will not soon so describe and control the totality of factors in any one student's growth that we can regard as unimportant the one which so much defines formal and informal education.

Although there have been some few studies which attempt to show a positive correlation between characteristics which identify effective teaching and specific student achievements, it cannot be claimed that convincing evidence of this kind exists. There is little evidence of the persisting effects of college instruction, whether learning measured in some respect at the conclusion of a course remains after a month or a year and what actual impact such learning has had. The area between teaching and learning is a large and still unexplored territory, one in which experience, enlightened common sense, and wise observation provide the best guidance.

Faculty distrust of observations by students often stresses the inability of students to do more than express their "opinion" of what they like and dislike about the teacher and his conduct of a course. Even if that were precisely true (and it is not true of well-constructed questionnaires), such responses would have considerable relevance to a student's learning and the professor's teaching. Learning theory still honors the long-established general notion that we tend to learn what we like; motivation is still a primary factor. The greatest contribution a teacher makes may not be in the subject matter he imparts but in his ability to incline students toward learning, in other words, toward "liking" what a teacher does. Teachers are not working with some abstractions of "learner" and "learning" but with the students who come before them. If a teacher registers strongly on the student's sensing apparatus, I think we are justified in connecting that with the student's learning. If a teacher fails to make an impact, we surely must have doubts about how much he is helping a student learn.

Purposes of course and teacher evaluations

These cautions raise the general question of the purposes of course and teacher evaluations whatever means are used, and the specific questions relating to the use of student opinions of faculty. Obviously course evaluations are useful to the students as guidance in selecting courses and to teachers in improving upon courses. Course guides which are primarily descriptive are hardly more than adjuncts to the college catalog. Indeed, the need for course guides may have developed as the size of a university's program reduced the amount of information conveyed in the catalog. But course guides naturally move to evaluation, for students basically want to find out whether this course is worth taking or is to be preferred to another. Courses have always been identified with teachers, and what may have begun as a guide to courses ends up as both a descriptive and evaluative instrument applying both to the course and the teacher. Any one of these instruments can provide the teacher with useful information. This discussion is primarily concerned with the use of instruments which furnish student response to the faculty member's teaching.

Evaluation of a teacher's performance has two closely related purposes (1) to render judgment on an individual teacher's effectiveness as a part of passing on his competence as a faculty member; and (2) to afford the teacher the means of developing his own skill. Although the one is commonly regarded as an institutional objective and the other a personal one, it is difficult and probably inadvisable to separate the two. Institutional judgment is often instrumental in developing individual skill, and conversely, the developing of individual skill serves the purpose for which institutional judgments are intended.

The case for better means of evaluation for institutional use is well made by the joint committee which proposed this project:

> The inadequacies of our present procedures for evaluating and recognizing teaching effectiveness are most apparent where decisions are being made about retention, promotion, and tenure. The decision for or against tenure is critically important both for the young teacher and the institution, but in thousands of instances each year this decision is made on the basis of extremely tenuous information about his teaching ability; a limited amount of student response, usually informal, undiscriminated, and distorted by successive reporters; judgments by deans, chairmen or colleagues deriving from social conversations and corridor exchanges; and guesses about classroom effectiveness based on the faculty member's performance of quite different institutional duties.

The committee's opinion rests on solid evidence. Clearly, the profession has a responsibility for making its most important personnel decisions on the basis of precise and searching criteria. In this sense, most institutions could improve upon the means of evaluating teaching and of evaluating all other aspects of faculty performance, including published work.

The use of evaluation in improving individual performance of teachers is no less important. Here the desired condition is to get away from pronouncing judgment and to use evaluation as we use criticism which we respect: to provide both the drive and direction for improving upon what we do. Outside observation is in general necessary to increasing the individual's purposeful insight. As such observation affects teaching at large, it can serve to make teaching a more open enterprise in which learning becomes more genuinely the common task of students and teachers to arrive at growth in knowledge, understanding, and perhaps even in wisdom. The members of the evaluation committee agreed that this purpose of evaluation was more important than the judgmental, and that an evaluation which restricted its function to the judgmental only was probably poorly employed.

Support for student evaluation

Assuming then that a college or university moves toward the evaluation of teaching with improvement uppermost in mind, what possibilities does student evaluation possess and what drawbacks? The following points are those which emerged most clearly from the many energetic discussions of evaluation in which I participated this year:

1. Student evaluation increases the chances that excellence in teaching will be recognized and rewarded. It may be the first step to a campus-wide consideration of the qualities that affect teaching and learning and to attracting into the profession those who see teaching as an attractive calling.

2. Student evaluation provides a means of participating between students and teacher which, in large schools, may not be present in other forms. Such participation may lead to the kind of involvement by student and teacher in the teaching-learning process which raises the whole level of instruction. Evaluation can give teaching a kind of dramatic interest which abstract discussion seldom affords.

3. It provides the only direct and extensive information about a faculty member's teaching. Such observations, unlike visitation, do not intrude into the class and are made by those with a genuine interest in the teacher's success. (It is no accident that campus-wide student evaluations tend to arrive at an average level of performance higher than the actual middle. Who wants to be the student of a below-average professor at a below-average college?)

4. An institution may be stimulated by student evaluation of teaching to consider its overall goals and values. If teaching itself is to be evaluated, then what value is teaching to have in relation to research, publishing, and service? If criteria for effective teaching are to be established, what about criteria for effective administration? And what about examining all parts of the rigid and traditional structure which fits individuals into the academic enterprise?

5. The support of student evaluation is a tangible sign that faculty and administration recognize the importance of student involvement in shaping the institution's educational goals and practices. In the words of the Princeton report, "The more we can persuade students to comment thoughtfully and responsibly on their courses and instructors, the more we can convince them that their comments are seriously considered, the more Princeton can do for its students and for the cause of higher education in this restless and rebellious age."

Common criticisms

The year's discussions also provoked some commonly expressed criticisms:

1. Student evaluation does little general good and some particular harm. Faculty anxieties on this point are not precisely clear, for the same critics who are most dubious about general impact are often very sensitive to the particular harms student ratings can inflict. Probably evaluation has impact, good and bad, in relation to the degree it moves away from the strictly private and personal. The effects of published evaluation are not so different from the effects of publishing books, speaking out in public, committing oneself in any way beyond the confines of one's own study.

2. *Student evaluation may arouse an unhealthy competition among faculty members. Such an argument is not likely to appeal to students who see themselves in an intensely competitive atmosphere forced upon them by the faculty. Although student evaluation may incline toward the "popularily contest," it can also have a salutary impact upon professorial vanity, arrogance, and pomposity.*

3. *Evaluation systems move toward harmful formalization, mechanization, rigidity. The relative merits of informal as against formal systems need to be continually examined in all matters of university functioning. Student evaluation in this respect is probably not much different from the operations of the registrar's office or of the teacher who seldom thinks past the 50-minute hour, the formal final exam, the bell-shaped curve.*

4. *Student opinion questionnaires furnish inadequate or misleading information about teachers and teaching. Objections of this kind are seldom based upon an acquaintance with the accumulated scholarly work which lies behind current questionnaires. Because of the existence of these resources and experience in using them, the creation of a valid questionnaire is probably the easiest task to be accomplished in instituting an effective student evaluation system.*

5. *The effect of student evaluation is a short term one connected with the novelty of the procedure. However true this may be, it also applies to most other procedures which once initiated become part of university routine. Institutions should have as part of their routine functioning the periodic examining of all routines.*

6. *Students by themselves will not be able to sustain a high level of evaluation procedures over a period of time. This is but another way of saying that if a college or university is to have student evaluation, cooperation and support must come from the faculty and administration.*

The above does not pretend to bring forth all the arguments pro and con on the subject nor to explain the nature of such arguments. My experience in working with evaluation this year, on campuses which have had student evaluation for a number of years and on those where it is being approached for the first time, indicates a high degree of receptivity to the idea, less resistance to the particulars than I expected, and a willingness to investigate practices and procedures. Adoption of student evaluation systems seems always to be a matter of convincing a good part of the faculty that such systems are useful and will not work direct harm. At the point where a college university is ready to make use of student evaluations, I think these are some of the major considerations:

A common enterprise

It might be well to begin by moving somewhat away from both terms: "student" and "evaluation." Something as neutral as "college or university survey of courses and teaching" might reduce friction and for all its generality better describe the system. It would recognize that while student response provides the basic data, the enterprise itself is a college or university enterprise, and not *for* the student or *against* the faculty or *by* the administration.

Cooperation among all three constituencies is vital to the success of student evaluation, both as a matter of instituting and administering evaluation procedures and of improving teaching by their use. In observing some evaluation procedures that work very well, I am inclined to say that the favorable conditions which surround the evaluation system are more important that the system itself. I have also observed these favorable conditions are fairly rare. Divisiveness is a characteristic of many campuses today. And although student evaluation could be a cause for greater divisiveness, I think it would more often work the other way. If students, faculty, and administrators become genuinely interested and involved in evaluating teaching, the campus may gain more from an improvement in the climate for learning than from the evaluation itself.

Other means of improving instruction

Such cooperation should also show itself in other measures taken to improve instruction. Discussions of evaluation procedures are discussions of teaching, and faculty members should make the most of the interest evaluation provokes among students. Responses can be many: more open doors in classrooms, more discussion of teaching; more attention to teachers learning from students as well as students from teachers; more useful attention to new teachers and their development; more help tendered, suggestions offered; in short, a better climate for teaching and learning.

In addition to these general effects, evaluation data from students may help move toward adequately recognizing and rewarding teaching. The willingness of a campus community to support student evaluation may be concrete evidence that the institution will support undergraduate teaching. As an evaluation gains the respect of the whole academic community, so can it become an actual part of the criteria by which the decisions on promotion, tenure, retention and salary are made. In time, a college or university might be able to show an actual positive relationship between advancement and demonstrably effective teaching, even though that demonstration might rest principally upon student response data.

It should be mentioned that though the majority of institutions follow similar practices in reviewing faculty performance, there are still schools in which practices are ill-advised or inadequate. The practice most dispiriting to teachers is that which imposes performance qualifications from the outside with little opportunity for the teachers themselves to help shape the instruments used or to have a vital part in the procedures. The performance check-lists brought into being solely by administrators are far from the kind of evaluation instruments being discussed here. Judgments which are made solely by supervisors cannot be expected to arouse enthusiasm or even cooperation among a faculty.

Finally, though I do not intend to envision Utopia, evaluation is related to increasing the individual's willingness to put energy and imagination into teaching and to enhancing teaching within the profession. The profession has argued for years that it cannot very precisely reward or even stimulate effective teaching without knowing what effective teaching is. Lacking that knowledge, it cannot very well identify good prospective teachers or attract them into the profession or help develop teaching skills. But we are not that short of useful information about teaching effectiveness, and the widespread use of student evaluations is generating more. We have little to lose in putting that information to work in examining what the teacher does and how what he does might be done better.

3. Student Evaluation Instruments and Procedures

Student opinion questionnaires, the most commonly used form of teacher and course evaluations, have a high degree of similarity. Such questionnaires grow out of examining other questionnaires or from exposure to such studies as those done at Purdue, University of Washington, the University of Michigan, and elsewhere. Such freewheeling publications as the Harvard *Confidential Guide* or Berkeley's *Slate* also furnish models, and work continues to be done on individual campuses in trying to improve upon existing forms. Current practice seems to favor the questionnaire which can be administered to large numbers of students in large numbers of classes and which results in a computer print-out of "scores" on the individual items. By now, it is quite possible to construct a questionnaire which presents the questions students and faculty members find important and in such form and with such weighting as to satisfy the minimum requirements of those professionally engaged in testing and evaluation.

Kinds of instruments

A number of practical distinctions can be made in describing the range of instruments used. First, there are those which use questions designed to bring out the nature of the course and/or its strengths and weaknesses as well as those which focus upon the teacher and his teaching. The combining of both kinds of questions on a single form is not uncommon, nor is it uncommon to find descriptive questions mixed with judgmental ones. Deciding on what kind of information is to be sought is an important matter, as is the careful construction of questions. Attention must be given to all the details of makeup which might affect response.

Another basic distinction is between short forms and long. The University of Washington questionnaire consists of ten carefully selected, placed, and phrased items. A form worked out by Wilbert McKeachie and published in the *AAUP Bulletin* (Winter 1969), contains 17 items on the teacher and teaching and six on the nature of the course. A very recent study at the University of California at Davis, based on analysis of 85 items that discriminate "best" from "worst" teachers as perceived by students, resulted in suggestions for questionnaires of various lengths. The shortest has fifteen items. The number of questions on questionnaires commonly in use appears to range from about ten to thirty, though I have seen one punch card form asking for only four responses and another form with over fifty items. In the light of the research upon which individual questions are based, brief, well-constructed questionnaires seem advisable for the general survey of opinion.

Questionnaires vary, of course, according to purpose. The questionnaire designed to gain information will probably differ from the one with a clear intention of providing data for judgment. An individual teacher may frame his own questionnaire to find out about specific aspects of a given course or his conduct of the course. Some departments create their own questionnaires, feeling that the nature of their courses and teaching is such that the broader questionnaire is not applicable. In general, the more multipurpose the questionnaire and the more specific the information desired, the longer a questionnaire will have to be.

Public and private

Another general distinction between questionnaires is between those intended for publication and those for private use. Questionnaires which can be machine-processed are most commonly used to provide data for published course or teacher surveys. The results appear as a profile of ratings — usually on a five-point scale — for individual courses and/or instructors. On such questionnaires, open-ended questions are limited though not necessarily ruled out. Students are often given the opportunity to make some short written responses to questions, though these may not appear in the printed results. The University of Texas course evaluation booklet prints a sampling of open-ended comments and a written response of the teacher to the student evaluation of the course. An intention to publish, like the necessity of reaching large numbers of students, has its effect upon the final form of a questionnaire. Neither places severe restrictions on the makeup of an adequate form.

Essay responses versus objective data

The debate over the validity and value of essay responses as against questionnaire data is difficult to resolve. It is not so different in character from the debate about essay against objective tests. It must be acknowledged that the widespread use of the objective test arises from the necessity of dealing with large numbers of students rather than from its proven superiority as an educational device. On the other hand, the motivating power of one as against the other may not vary greatly, and the actual outcomes from either kind of test may not be far apart.

The growing use of the computer-processed questionnaire is related to the size of our colleges and universities and to a necessity to provide interchange where personal exchange may have become difficult. Personal interviews with students about the nature of their courses and instruction take time and pose difficulties of their own. And though ideally we may argue that such personal attention is worth it, the practical question is whether faculty members will be willing to give it. There is a further question that useful information about courses or a teacher will be forthcoming from a face-to-face talk between student and professor. Such bluntness and honesty as may be required ask more of both parties than may be reasonable to expect. These practical considerations do not dispose of the need to retain an individual, personal dimension in evaluation forms and systems. Many "objective" evaluation questionnaires provide space for "subjective" comments. Such comments, offered anonymously by the student and returned only to the professor, may give a two-way purpose to formal and informal counseling.

Implications in the use of questionnaires

Finally, there are implications in the use of student questionnaires that may easily escape the faculty. Many of the questions raised about the use of questionnaires are really about the quality of the instruments used. There are good and bad questionnaires, and faculty members exist on most campuses who can ask the technical questions and provide the professional assistance to arrive at good ones. The faculty's professional competence has been drawn upon in shaping many student questionnaires. It has probably been less employed in the faculty's examination of its own criteria for judging an individual professor's performance.

No student groups of which I am aware have argued for exclusive right to evaluate either teachers or courses. On the other hand, the existence of a responsible student evaluation in itself argues for its inclusion among the criteria for judging teaching performance. At the University of California at Davis, the student Academic Affairs Council proposed that students and faculty make independent judgments of a professor's strengths and weaknesses based upon a variety of inputs. Though the proposal has not been accepted, its intention is to give equal weight to the faculty's and student's recommendations. In institutions where formal student input is actually being used, some combination of faculty, department chairman, and dean still has the major power in setting criteria as well as in judging performance. There is much to be learned from the research and experience in the broad area of student evaluation and probably much that could be applied to the faculty's own evaluative policies and practices.

Administering a questionnaire

Administering opinion questionnaires raises another group of important questions about student evaluation. What was once largely a student activity has become a somewhat institutional practice. The freewheeling course evaluation booklet compiled by free spirits without benefit of professional advice or institutional support often had the virtues of imagination, candor, wit, and bite. Even in these times, there is some point in emphasizing that fact. Such critiques have not flourished on many campuses probably because the writing and publishing of such critiques is a time-consuming extracurricular job. The administration of many current student evaluations is an institutional enterprise proceeding through a formal organization, a student-faculty-administration group, or an office of evaluation and testing.

Objectivity

The major considerations in administering a questionnaire are those of objectivity, coverage, and timing. The makeup of the questionnaire affects its objectivity just as does the way it is distributed and collected. Involvement of student, faculty, and administration in the making of questionnaires can be a way of eliminating some slanting that can come into the choice and posing of the questions. In the actual administration of the questionnaires, some distance should be preserved between the faculty member whose course or courses are being evaluated and the students doing the evaluating. Filling out questionnaires under the eyes of the faculty member who has handed them out and plans to take them up, though it may testify to a highly desirable degree of trust between faculty member and student, may also inhibit or distort the response. A designated time apart from regular class hours, a third party in charge of giving out and collecting forms, assurances that the student's anonymity is protected — these appear to be minimum necessities for gaining a measure of objectivity.

Coverage

The purposes and uses of an evaluation will also influence its manner of administration. Is an entire campus going to undergo systematic evaluation? Is the freshman program what most needs looking at? Are individual departments going to examine their courses and to what ends? Is the data going to be published? Is the data going to be used by anyone other than the individual professor? The size of the enrollment on a large university campus has a bearing upon the kind of instrument used and the coverage that can be attempted. Campus-wide surveys have a number of advantages: they involve the entire campus in looking at teaching; they can justify greater care and expense in preparing questionnaires and in examining data; they afford comparisons across departments, across the campus. At the same time, the survey which attempts to obtain usable data from all courses is a large undertaking. If such attempts are made, the survey might be extended through an academic year, with portions of it confined to one specific semester or quarter. Repeating surveys quarter after quarter or even semester after semester is probably less advisable than obtaining reasonably consistent data over a number of years.

The extent to which an evaluation survey covers the entire academic program is also a matter of how many students respond to the questionnaires. Obviously, a high percentage of response is desirable. In the very carefully administered Princeton surveys of the past two years, a 95 percent response, university-wide, was obtained during the first semester. The percentage dropped to 78% in the next term and 76% the following term. This percent of response still seems adequate to the committee at Princeton, but it is concerned to see if the percentage will drop or level off in future years. The questionnaire is given during the final examination period with a special portion of time set aside for filling out the form in class. Some campus surveys adopt what seems to be sound practice in only printing results based on 70 percent or higher response. However, data from all courses regardless of the percentage of response is made available to individual faculty members. In this connection, small classes, those under a dozen students to suggest a figure, may be inappropriate to include in a general survey. With such small numbers, the possibility that the nature of the course will provide adequate student response may obviate the use of a questionnaire.

Timing

A good deal of attention has been given to the best time to conduct an evaluation survey. Very close to or at the end of a quarter or semester is the favored time, although there has been some use of questionnaires deliberately designed to gather information in mid-course. The purpose of such interim feedback is to offer the opportunity for improvement while the course is still going on. The effect of grades upon responses to a course or teacher evaluation has been studied as much as any single aspect of the subject. Little evidence emerges to support a belief that grades determine response. It appears possible to use a questionnaire at the end of a course either before grades are given out or after. Research is not very helpful in deciding whether it is better to make a survey of a class while it is very fresh in the person's mind or to wait until some weeks or months have passed before asking for such data. Studies that have been made to compare responses about courses and instructors at the time with those made at a later period suggest a consistency of response. There may be good reasons, however, for sampling opinions from certain groups, freshmen for example on introductory courses, or majors on aspects of their specific program, or seniors looking back at their careers. The Princeton pilot studies in student evaluation include four questionnaires: the first to students in every undergraduate class, a second to all sophomores to evaluate the experiences of the first two years, a third to seniors in the senior year assessing their work in the departments, and a final questionnaire to be given to graduates six to twelve months after graduation inquiring about their entire educational experience at Princeton. Gathering and interpreting information from alumni poses more difficulties than conducting an on-campus survey, but colleges could probably profit from trying to gather such information. Again, such research studies as have been made suggest that alumni opinion about teachers and teaching is reasonably consistent with student opinion and consistent also with opinions the alumni themselves held when they were students.

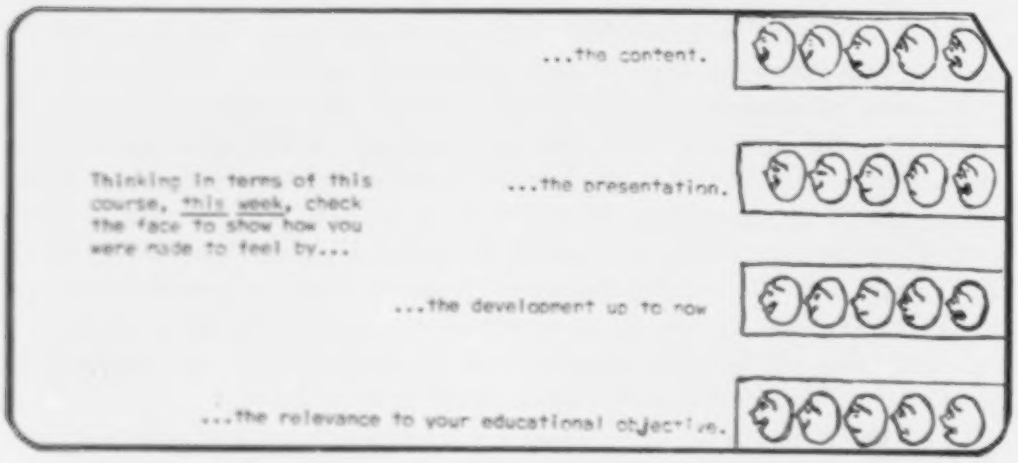
Constructing a questionnaire

The kinds of questions students can best answer are those which fall directly within their powers of observation and insight. The ten questions on the University of Washington questionnaire are all of this kind. Some: "gets me interested in his subject," "has helped broaden my interests," "has given me new viewpoints or appreciations," are particularly ascertainable by the student. Others are only slightly less so: "has increased my skills in thinking," "has motivated me to do my best work," for there may be some slight question that a student can accurately measure his skills in thinking or his degrees of motivation. The other five make judgments which are based on the individual's perception of specific teaching objectives: "is clear and understandable in his explanations," "makes good use of examples and illustrations," "interprets abstract ideas and theories clearly," "stresses important material," and "inspires class confidence in his subject." I have listed these in the order in which I think the student judgment is most reliable. Though what is "clear and understandable" to one may not be to another, the response of an entire class on this question is a very accurate measure of the degree to which the teacher is getting through. The same may be said of the question "stresses important material," although it may be argued that students judge what is important by what the teacher stresses. It should be kept in mind that students do not operate in a vacuum. They read books; they attend other classes; they discuss courses and teachers, assignments, tests, even the teacher's jokes, with other students. Even in matters where the students seem least capable of judgment, their assessment of a faculty member's grasp of his subject, for example, they are probably not completely in the dark. Communicating of a subject has something to do with knowing it, just as bluster and verbiage and evasion raise legitimate questions about a teacher's competence.

Questions on many questionnaires are not as carefully selected as those used at the University of Washington, though many of such questions are extensions or variations of these basic ones. Some questions deal with observable aspects of the teacher's behavior — "has a sense of humor" — which may not be crucial to teaching effectiveness but is more likely than not to be a positive trait. Negative characteristics — "frequently employs sarcasm" — are also included. The range is fairly wide for questions which still focus on observable acts of a teacher's behavior about which the combined responses of students are likely to give an accurate picture. In many ways this is the kind of feedback an individual could get from observing himself in action by using an audio- or video-tape. Making a useful self-analysis, however, may be more difficult than working with student responses. Many questionnaires include questions about relations between students and teacher which are both descriptive and which involve judgment. "Respects student's opinions," "friendly toward students," "readily available outside of class," are examples. Another group of questions concerns the handling of instructional details, e.g. "returns assigned work promptly," "fair and impartial in grading," "assigns appropriate materials," "gives fair tests."

As one moves away from those questions which are the result of careful analysis based on data from large groups of students over an extended period, the connection between observed performance and actual effectiveness probably becomes more tenuous. A good questionnaire places a desire for maximum amount and specification of data against that minimum which has highest reliability. My own conclusion in this matter is that a relatively short questionnaire like that in use at the University of Washington is an excellent model for general, campus-wide use. Wider-ranging questionnaires could be employed to gain more specific details of performance or student response. I personally would not rule out many questions that stray from those most closely supported by studies of relevant criteria. To a great degree, student questionnaires furnish answers to the questions we ask. If we think it important that a teacher show up for class on time, students can tell us if he does. If we believe arrogance toward students has a bearing on effective teaching, the opinions of many students over a range of courses provide a means of identifying this attribute. If we value imagination and innovation in instruction, questions can be devised to emphasize those values.

The "we" I am referring to are the members of any campus community. For the whole community, in some workable representative way, needs to be involved in shaping, testing, and using student evaluation materials and procedures. At the same time, such cooperation should not serve to detract from the particular competence, point of view, or enthusiasm of one group or another. It is cheering to know that the University of Texas physics department makes full use of a questionnaire made out by graduate students around a tavern table; that some questionnaires, despite the horror this arouses among testing officers, use a scale of ABCD ratings; that a quick feed-back card used successfully at the University of Wisconsin-Green Bay appears in this form:



Thinking in terms of this course, this week, check the face to show how you were made to feel by...

...the content.

...the presentation.

...the development up to now

...the relevance to your educational objective.

Establishing and maintaining an evaluation program

Most evaluation efforts are initiated by students, either as an independent venture or as an activity which seeks official university recognition. Both have their place, but I am going to discuss here only the establishing of those systems which seek to gain college or university acceptance and use. The following is a list of necessary actions and accomplishments if a successful student evaluation of this kind is to be established and maintained:

1. *Gaining the cooperation of the faculty.*
2. *Defining purposes, objectives, and uses.*
3. *Arriving at means and procedures.*
4. *Making crucial policy decisions.*
5. *Establishing an office for administering the program.*
6. *Keeping the campus community informed.*
7. *Financing a continuing program.*
8. *Maintaining student and faculty interest and involvement.*
9. *Conducting follow-up activities and studies.*
10. *Relating evaluation to other efforts to recognize, reward, and improve teaching.*

Each of the above items could be discussed at length, but I will attempt only to touch upon those matters in each category which seem to be most important.

1. *Gaining the cooperation of the faculty.* Student evaluation has had an evolutionary growth and acceptance. When the idea of conducting an evaluation came up at the University of Utah in the late fifties, it was summarily rejected by the Faculty Council. A few years later, the Council was willing to study the matter before deciding that proper evaluation procedures would cost too much. Five years after that, students were able to gain sufficient individual faculty support to develop and administer a campus-wide questionnaire. Faculty members, it seems to me, can be reached by the rational arguments upon which student evaluations are based. What the faculty wants of student evaluation procedures is not far different from what students want of grades: (1) a reasonable basis for making judgments; (2) fair and responsible assessment; (3) a clear idea of the purposes, objectives, and uses to be made of evaluation instruments and data.

2. *Defining purposes, objectives, and uses.* Opportunities for members of the campus community to discuss the issues raised by student evaluation are vital to gaining support. Such discussions will go a long way toward arriving at the kinds of instruments to be used and procedures to be followed. Discussions should begin very early, for then is the time to clarify common aims, to reduce suspicions, and to view the subject in as objective a way as possible. On the basis of my visiting campuses this past year, I think the presence of a knowledgeable outsider can be very helpful. The visitor's expertise aside, his detachment from local campus affiliations, factions, or

past controversies may help give the idea of evaluation a reasonable hearing.

3. *Arriving at means and procedures.* One of the benefits of early campus-wide discussions is the identification of students, faculty, and administrators with special interests or competence in various aspects of evaluation. Cooperative departments may also be identified, for initiating an evaluation program in a single department may be a good way of shaping and testing instruments and procedures. Princeton's evaluation system was first tried in four departments before gaining campus-wide adoption.

To me, however, the most important single point to be made about adopting instruments and procedures is the necessity for colleges and universities to make use of the accumulated research and experience with student evaluation. The existing literature provides a wealth of materials — most of it readily available for any school's use — for constructing various kinds of student response questionnaires. Energies can be best spent in considering how tested materials and experience can be adapted to a specific campus situation.

4. *Making crucial policy decisions.* Too often, student evaluations come into use on the basis of stu[illegible] initiation and vague faculty assent. The faculty accepts evaluation as a [illegible]student activity" and avoids facing the difficult questions that an effectively-operating evaluation system is sure to raise. Some of these questions are:

1. What members of the faculty are to be evaluated?
2. How often are evaluations to take place?
3. How are the results to be made known?
4. What use is to be made of the results?
5. Are student ratings to be given official recognition as part of the criteria for decisions on retention, promotion, salary, and tenure?
6. What weight will they carry?
7. What other ways will be used to ascertain teaching effectiveness?
8. Will students have other opportunities to be heard in matters affecting teaching and learning?

These are important policy questions about which clear decisions need to be made by the established representative campus governing body.

5. *Establishing an office for administering the evaluation program.* Someone has to do the work of conducting an evaluation. Students are most often charged with the task, but students are a transient population, and the quality of student surveys too often depends on the competence and in-

terest of students currently involved. The student affairs office, a specially created student-faculty committee, the office of evaluation and testing, are some of the administrative groups used to give greater consistency and continuity to student evaluation activities. My own preference is for the establishment of a center specifically charged with the promotion of effective teaching and responsible to the faculty senate or other university-wide governing body. It would be staffed by both students and faculty, and its activities in evaluation would be only a part of its interests and activities. Ideally, prominent faculty members would serve full-time as members of the center on a semester or annual basis. It would have attached to it a professional staff member skilled in such matters as learning theory, teaching techniques, course development, evaluation and testing. It might well have as its guiding purpose the initiating, encouraging, and supporting of individual and departmental efforts to improve teaching.

6. *Keeping the campus community informed.* I have stressed the importance of involving faculty, students, and administrators in shaping the purposes, the instruments, the procedures, and uses of student evaluation programs. More than that, continuing opportunities need to be provided for members of the constituency to know what is going on as the program is developed and used. A basic part of establishing a questionnaire should be the maintenance of easy means by which individuals can register objections and make suggestions. Another part should be provisions not only for review, but for making the substance of that review known to members of the campus community and again giving individuals the opportunity to have a part in revising and further shaping.

7. *Financing a continuing program.* The precise cost of an evaluation will depend upon how inclusive it is to be, what kind or kinds of questionnaires are to be used, whether or not the results are to be published, and the like. In administering most questionnaires, the use of student manpower and institutional facilities cuts down on the actual dollar cost. At Princeton, the bare cost of administering the basic questionnaire which went to all students and resulted in a computer print-out of results was estimated at $1 per student. The Academic Affairs Council at UC-Davis proposed a budget of $5,000 to cover costs of reviewing the faculty, with priority given to professors about to receive tenure. At Michigan State, the first year's course booklet in 1967-68 cost approximately $2500 for surveying about 180 classes. This included a very minimum for secretarial services, questionnaires and IBM response sheets, computer operation time, publicity in selling course booklets, and printing costs. A well-done printed evaluation booklet can often

be sold in sufficient quantities and at a price which will offset a large part of the costs. None of the costs suggested above include the price of developing a responsible questionnaire. However, the costs here need not be great if an institution is willing to draw upon work already accomplished and available.

The main point about financing is a simple one. The costs are not such that an institution could not find funds if it thought the activities involved in student evaluation were worthwhile. If financing is left for the students to provide, or if a printed booklet is expected to make this operation pay for itself, the results may well suffer from the inadequacy of support. Student evaluation aimed at the recognition and improvement of teaching is a college or university activity and deserves direct support from the institutional budget.

8. *Maintaining student and faculty interest and involvement.* I do not think institutions stand to gain very much from student evaluation unless they make provisions that it will be well done and that it will operate sufficiently long to make it possible to see results. In addition to the necessary attention to conducting and maintaining the quality of the evaluation itself, attention needs to be given to maintaining student and faculty interest and involvement and to the other matters on this list: *conducting follow-up activities,* and *relating evaluation to other efforts to recognize, reward, and improve teaching.* At least two follow-up activities are important: (1) exploring ways that evaluation data can be of use to members of the campus community, and (2) conducting studies and activities that examine the effectiveness of the program. The last function would be served in the establishment of the campus center for teaching previously suggested. Once again, it is pertinent to repeat the committee's firm belief that evaluation is only one of many efforts which might be made in support of teaching. Student evaluation needs always to be viewed within a larger context of attention to teaching, and specific development of an evaluation program as part of a wider effort to enhance learning.

Other questions

Materials in the Appendix discuss in more detail many of the questions that are raised about student evaluation forms and procedures. Certainly data can be defective and the procedures poor. But such defects are not inherent in the forms nor is it very difficult to arrive at satisfactory procedures. Student questionnaires can only partially if at all furnish accurate measures of a teacher's competence in his subject matter, or the actual worth of the material assigned or presented, or the degree to which his teachings are grasped and retained. The first two are those which his peers should be willing to find out about and be able to judge. The last is a question to which the combined efforts of students, faculty, and administrators should be seeking answers.

4. Impact of Student Evaluation

Student ratings which are seen only by the instructor have been widely accepted as a means of improving teaching. Private responses are useful to help the teacher improve his teaching, to examine and improve upon the details of a course, to give students a general sense of participation, and to increase the student's own thinking about teaching and learning. Studies which attempt to examine the effect of such private feedback rely greatly upon subjective comments from the teacher. Comments vary from noting very little effect to acknowledging considerable assistance in specific matters of classroom conduct and in the development of general teaching skills.

The need for feedback

The skillful teacher, like the skillful actor, may not need the feedback such questionnaires can provide. He may already possess a keen and constant sense of the relationship between what he is doing and the way the students respond. In addition, the teacher has other measures — tests, papers, oral reports — which give him information about the way a course is going. But teachers must acquire teaching skill, and even skilled teachers fall into comfortable routines. The audience is seldom the actor's audience and professional criticism of teaching is not a part of academic tradition. The analogy with acting risks arousing the profession's distrust of the teacher as performer, and pursuing that analogy is not my intent. The comparison serves rather to emphasize the necessity of feedback of some continuing kind if the teacher is to make an impact upon the students and to develop and improve upon his teaching skill.

However helpful private student responses may be to the individual teacher and his courses, they are a limited way of improving teaching. At the least, such activity needs the kind of support that makes teaching a matter of importance for the whole campus as well as for the individual faculty member. There is little doubt that the initial appearance of published evaluations on a campus creates a dramatic interest in teaching not often provoked by other means. Faculty members are as sensitive to public opinion as other human beings, and open student evaluation has this force. Even when there is a reasonably clear understanding that evaluation data will not be officially used, the presence of a responsible published evaluation is a constant reminder of the necessity to teach well.

Impact of open evaluations

Open evaluations do not need to be published to have a greater impact than private ones. Once an evaluation becomes open to the scrutiny of others — department chairmen, deans, or, if it is published, the whole university — the results have consequence not only for improving teaching but for passing judgment on a faculty member's work. Just as it is difficult to separate the descriptive from the judgmental in the items on a student questionnaire, so is it difficult to separate information from evaluation in the uses made of student responses which are not held in confidence. The formal judgments about a colleague's competence as a teacher are made within the present system upon the basis of whatever the faculty members know or can guess about that colleague's teaching. If published evaluations are abroad, faculty members will read them. If evaluation results are officially made known to the department chairman or dean, it is fairly certain they also will use them.

As student evaluation responses become other than private information for the benefit of the faculty member, their possible effects come not only from the advice furnished the individual professor but from the impact of wider dissemination. These effects must be carefully considered, for they operate whether or not evaluation is incorporated into stated institutional procedures. Certainly it is important to reduce the anxieties that open student evaluations arouse. To an insecure teacher with or without tenure, the potential threat to his own self-esteem is probably as great as to his position on the faculty. An adverse student rating can destroy confidence and worsen the postures of withdrawal and defensiveness which are destructive to effective teaching. Academic tradition supports an atmosphere of freedom in which surveillance of any kind is often resented. Is a faculty member who is already expected to perform well as a scholar, teacher, and contributor of services likely to respond positively to pressures to teach better? Is he not likely to resent evaluation systems which increase that pressure?

There is a basic argument that teaching is of such importance and the general quality so low that strong pressures need to be exerted. The image of a professor crushed by adverse student opinion is opposed by that of students forced to suffer the incompetence of an untouchable professor. On some campuses, this argument may be valid. Within the profession at large, the presence of both images has to be acknowledged. The real or imagined threat in student ratings can be reduced, needs to be if they are to offer a continuing means of improving instruction.

Reducing the threat

A number of common-sense measures come to mind. First, faculty members must be informed and involved in the adoption of an evaluation system. Second, they must be given continuing accurate and detailed information about the use of the results. Third, ways should be clearly established for individual faculty members to register objections, to particularize complaints, and to have assurance that instruments and procedures are responsive to criticism and suggestions. Fourth, opportunities need be freely given for private counseling about an individual's teaching. Similarly, forums for discussion are important not only to making the most of the interest which evaluations arouse but to establishing confidence in the system through the free exchange of actual experiences with it. These measures seem to be better than attempting to set restrictions upon the uses of open evaluations. Such restrictions tend to create new tensions over whether, in fact, the results are being used. The direction of all the above suggestions is to encourage an openness about teaching, a candor with oneself and with one's colleagues, and a willingness to let teaching be discussed and made better from the discussion.

Several cautions should be added. Beginning faculty members are particularly vulnerable to the effects of adverse student ratings. (It is also my observation that they are more receptive to the idea.) On balance, I think it is probably not wise to exclude them from a general open evaluation; they benefit from the feedback and from the comparison of their efforts and methods with others. What is important is a department's awareness of their inexperience and its willingness to provide counsel and assistance. Such direct concern can be augmented by attention to tangible gains that can come out of evaluation, such as sharing of ideas and methods between the new instructor and the established ones, and by recognizing not only excellence in teaching itself but the distinctive characteristics of individual teachers. The need to provide a fair and reasonable basis for the choices that must be made in developing a strong teaching faculty is reason enough for emphasizing the non-threatening aspects of student evaluation. The tensions evaluation creates may not be as serious as those growing out of personnel decisions which seem to be based on even less substantial data.

A poor teacher made better, a teacher reassigned to make the most of his kind of competence, a superior teacher contributing to the development of other teachers, an atmosphere in which teaching is worth doing well — these are some of the positive aspects of using student evaluations. Teaching is such a personal relationship that it poses its own constant threat in the class which does not respond with some favor to what the teacher is doing. Every teacher experiences some classes like this, classes which get started wrong or which push into a blind alley where students and teacher are genuinely at odds. Coming to such classes is a daily trial. In most instances the classes never right themselves and might better be called off than continued. The impact of student ratings can only approximate this kind of painful confrontation. In furnishing guidance to students and faculty, ratings may reduce this kind of threat already present in teaching.

Popularity Contests

The charge that student evaluations are merely popularity contests and sources of unhealthy competition is somewhat the mirror image of the first argument. Here the claim is not that individual teachers are being excessively threatened by adverse criticism, but that the faculty as a whole is being hurt by excessive praise going to the popular professors. It is difficult not to be impatient with this argument, for behind it lurks the grim schoolmaster who will measure the worth of instruction by the pain it inflicts.

Various studies of student ratings suggest that popularity in its meretricious sense is not what students are after in asking for better teaching. The commonly used questionnaires do not emphasize this kind of popularity, and an examination of the essay responses students make supports the belief that students are adept at recognizing that kind of popular professing. Nor do the studies confirm the suspicion that professors can win high ratings by easy grading and light assignments. "Tough" teachers are commonly highly rated. Their popularity can be measured by the numbers of students wanting to get into their classes.

But these observations aside, questionnaires can be designed to reflect the values prized by the faculty as well as by the student. Ratings, therefore, are measures not of abstract "popularity" but of specific characteristics students and faculty have used to define effective teaching. Although the halo effect has not been particularly noticeable in studies of student evaluation, it is likely that faculty reputations may be reflected in ratings as in student scuttlebutt. However, the aim of student evaluation is not that of singling out the popular professor or the favorite professor or even the most distinguished professor. It is rather that of gaining information about a great number of courses and professors and making it available for the guidance of both teachers and students.

Student evaluations do not aim for some single point score by which Professor A is judged to be two points better than Professor B. The results from itemized questionnaires are most frequently presented in some form of profile which is both a more informative and a wiser way of presenting results than through a single numerical average. Since each faculty member differs in strengths and weaknesses, exact comparisons cannot be made at a glance, through profiles of departmental or university norms permit general comparisons.

Emphasizing the necessity of teaching well

There is an element of competition in the use of student ratings which cannot be eliminated. Yet in asking about student evaluation systems which have been in operation for a number of years. I have concluded that the competitive aspects diminish rapidly. Pressure to excel as a teacher is not likely to be as intense as pressures to publish even with student evaluations as a part of that pressure. Admittedly, if the necessity to teach well were emphasized throughout the profession, the lot of those who neither taught well nor published satisfactorily would become more uncomfortable. But even then, such an emphasis upon teaching would widen the range in which a faculty member could develop and demonstrate competence. The argument against student evaluation as a popularity contest has not carefully considered the nature of student evaluation instruments nor the actual uses to which student evaluations are put. Although student evaluations can be used as part of the basis for selecting outstanding teachers, they exist on many campuses which give no such awards. In very few instances are student evaluation systems created for that purpose, and I have never heard of a listing of professors in order of their ratings on a student evaluation.

Student evaluations, because they are extensive, because they deal with specific characteristics, because they are created chiefly out of those matters which students perceive as helpful to their learning, tend to describe teaching even as they exercise judgment. If a college or university wanted to emphasize competition, it could use student evaluations to do so. If it wanted to banish dullness from the classrooms, evaluations might even assist in that.

But the competitive and purposeful aspects need not be given emphasis. Student evaluation can be useful simply in the visibility it gives to teaching and in the opportunity it provides for student response. Faculty members should have a strong voice in determining how student evaluations will be used and in seeing that their use contributes to the kind of campus climate desired.

Encouraging fresh teaching

One last general criticism remains to be discussed. This is that the use of open evaluations will favor the safe teacher and discourage the innovative one. Once again, much depends on the questions asked. If the questionnaire emphasizes the experimental and innovative then it may be a way of encouraging both faculty and students in this direction. If the campus climate is favorable to experimentation and innovation, it will reinforce the effects of such a questionnaire. The appearance of student evaluations on a campus for the first time is in itself an indication of receptivity to what is still a new procedure. And though a teacher might elect to play safe and gain modest approval rather than risk more and experience public failure, I suspect such choices are made with or without an evaluation system.

The chances that student evaluation will encourage innovative teaching are at least as good as that it will discourage it. Professors who read from yellowed notes are as much an academic fact as folklore. The teacher who has ceased to learn, ceased to place his past learning in relation to changing events, falls short by most standards of what a good teacher should be. It is easier to teach an old course in an old way than to try something new, just as it is easier not to read a book than to read one, and easier to read than to write one. For these reasons I think the energy expended in worrying about the inhibiting effects of evaluation could be better used in employing evaluation as an agent in stimulating innovation. One of the most beneficial effects of research into student evaluation of faculty is the identification of a variety of teaching styles. Evaluation procedures can be designed not only to be hospitable to a wide variety of techniques, but to encourage the development of new teaching skills by identifying specific teaching characteristics.

Maintaining student interest

There is a final aspect of using student evaluations which faculty members rarely consider. That is its impact upon the students. If a student evaluation is to be used, the faculty should be concerned not only with the quality of the instruments and the procedures followed, but with maintaining widespread student interest and a high level of responsible participation. During a time when student activism dominates higher education, it should not be forgotten that great numbers of students are still not being consequentially affected by their college experience. Colleges and universities vary widely in the character of their undergraduate students, but few large colleges and universities have enough students actively involved in their own education. So it is that student evaluation programs founder as much from lack of student involvement as from faculty opposition. On the other hand, the impact of evaluation may be fully as great upon the apathy of students as upon that of professors. Insofar as evaluation serves an educational purpose — and there is no defense for it if it does not — the faculty need concern itself with encouraging students in this and other endeavors aimed at increasing their responsiveness to teaching and learning.

Legal implications

A mention should be made of the propriety and legal advisability of using student questionnaires. Part of the impact of published ratings is the sense of injury adverse ratings can create. The question of legal action is raised infrequently, and the literature reveals very little in the way of specific charges, cases, or adjudication involving student evaluations. Faculty members have the protection accorded all citizens by the laws against libel and slander, and such laws could be expected to apply in the case of published evaluations as in other published expressions of opinion. Publication raises the possibility of unintentional injury that erroneous information can inflict. But such a possibility is not offered as a serious argument against a free press, and the established policies of correction and apology should serve for the limited circulation of most published evaluation booklets. If strong sentiments against the propriety of evaluation or murmurings about legal actions arise on a campus the climate is probably not such that evaluation would do any good. I cannot believe that teaching with its need for trust, openness, and freedom of expression could thrive in such an atmosphere. Nor do I think the impact of student evaluation would be sufficient to change it.

Comeback Professor of the Year

I will end this discussion with a story about a specific instance of the impact of a published evaluation on a faculty member. In spring of 1968, a student review of courses and teachers found a required senior course taught by a very reputable professor seriously deficient. Classes were described as being invariably lectures which followed the textbook almost line by line. The teacher's presentation was "slow, confused and aimed generally at the blackboard in tones that are barely audible." He was charged with being unprepared and not interested in the students or his teaching. The review concluded in this way: "Careful consideration is given before a write-up such as the one you have just read is published. Though not quite half of the total enrollment was polled, the results were extremely uniform. This course has been presented much better in the past, though this is the first time *Confidential Guide* has reviewed it. . . . Many students consider this course the worst they have taken."

The fall 1968 course guide noted improvement over the spring, but still called attention to the professor's trouble in communicating. His overall teaching ability was rated as "pretty close to average," and students found his outside help easily available and very worthwhile. By spring, the professor's deficiencies in teaching a class were still being noted, but his interest in the students and help given them outside class had gained favorable attention. In fall 1969 the student guide referred to him as "official winner of the Comeback of the Year Award." The classes were still basically lectures, but time was set aside to discuss problems. Class attendance was still optional, but "as the classes are generally worthwhile, few miss many." The professor was described as being exceptionally competent and a fine instructor. He had turned into a good speaker, was interested in his students and available for outside help. Tests were tough but graded fairly and moderately. "This is basically a hard course because of the subject matter," the review concluded, "but Professor ——— seems to have been able to make it easier on everyone."

5. Student Evaluation and Faculty Review

Considering the variety of teachers and students and their diversity of talent, training, and inclinations, the profession should be receptive to all those means by which a teacher may gain insight into his courses and manner of teaching. Though some faculty members scoff at the notion that student questionnaires do any good, few flatly oppose their use if the results are made known only to the faculty member. Objections arise somewhat in proportion to the degree that evaluation data becomes public and seem likely to have an impact. A good many colleges and universities which have student evaluation systems disclose the results to the department chairman or dean as well as to the individual faculty member. This may be required by institutional policy or left to the choice of the faculty member. A large number of schools have moved to open evaluations either through published booklets or by making data available in a central location. But not many colleges and universities have given student evaluation data a clear and explicit place in the actual making of decisions on retention, rank, salaries, and tenure. Within the last year, two large institutions, San Diego State College and Washington University in St. Louis, made widely-publicized announcements that beginning with the new academic year student evaluations would be added to the criteria used in making decisions on faculty advancement.

Student evaluation and faculty advancement

The conference on evaluation unanimously agreed that student evaluations should be used in making these important decisions. That consensus was based, I believe, on a recognition that students could contribute valuable information about teaching, and that such information should be used in an effective way. Although the group did not arrive at any specific procedures for using student input, it clearly was thinking of more than token participation. The direction of the committee's thinking was that if student evaluation were to be done, it should be done well. If it were done well then it would yield data which was highly significant in evaluating teaching and which existing procedures did not provide.

The simplest official use of the results of student evaluations would be to incorporate them into the faculty review procedures. Ordinarily, departments set the specific criteria for judging a faculty member's performance and have the responsibility for formal review of the quality of performance. These are not, however, exclusive responsibilities of the department, for the college and university play a part in formulating criteria as well as in carrying out review and reaching decisions following the department's recommendation. Although some institutions perform searching annual reviews of each faculty member, the most rigorous scrutiny is made at the time of decisions on retention, promotion, and the achievement of tenure. At these times, department recommendations must usually go to the dean of a college for further review and recommendation and perhaps to a university-wide committee for additional review before being submitted to the administration for final decision. In some small colleges, a college-wide personnel committee has major responsibility for review in all departments. Few institutions leave these decisions to a single authority, department head, dean, or president. In almost all instances, the colleagues nearest to the faculty member undergoing review play an important part in the recommendation.

Student input could be added to this generally-applicable procedure at a number of points. It could be used at the department level where the most knowing but perhaps the least objective review is made, or at any of the other review levels, or at all of these. It could take the form of data furnished to those in charge of the various reviews, or it could be treated as separate data which accompanies the formal written recommendations. It could be a part of a formal review by students within a department's review procedures either by use of a student committee or by participation of students on the departmental review committee. Such forms of active student participation could also be a part of the procedures at any other level of review.

Current faculty review practices

But before discussing some of these possibilities and as a means of informing that discussion, let us consider some aspects of current practices.

First, the present system is one which stresses judgment more than development. The reward and punishment motivation which colors all education seems to have its effect upon basic personnel policies. The annual review is for many younger faculty members something like the deliberations of a Ph.D. orals committee. The examination, however, has not just been given; it has been going on all year, or if tenure is at stake, for four or five or seven years. Worse for his peace of mind, the expectations may not have been clearly stated nor indications of how well he was fulfilling them very clearly given. Many reviews take the form of elaborate musical chairs. Since discussions are confidential and only members of a higher rank can make decisions on the lower ranks the group gets smaller and smaller until only the tenured full professors are left. The atmosphere is not one of mutual assistance toward common goals, but of judgment and sentence, or, in the common academic tongue, "up or out." Nor are the years previous to decisions very often years in which faculty development has been a conspicuous part of departmental or university practices.

The second aspect of review policies is the secrecy in which decisions are reached. Although, as with all aspects of these matters, practices vary greatly, some secrecy or privacy is probably necessary to all practices. The actual decisions are probably never reached in the presence of the faculty member, if only because actual decisions are made up the administrative line. Department chairmen may talk at length with the individual faculty member and even indicate that a positive or negative recommendation is forthcoming. But the crucial discussion in which a departmental chairman's recommendation is placed before a department seldom, if ever, takes place in the presence of the faculty member under consideration. I am not saying it should or bothering to rehearse the obvious reasons for confidentiality in these proceedings. One does not have to pass on the merits of such proceedings to observe that such actions create tensions and emphasize the necessity of providing for other communication between the individual and the group.

Traditional ways of handling these matters are currently under pressure from at least two sources: the students demanding an active voice in the process, and collective bargaining with its linkage to labor-management kinds of decision-making. Both pressures work against the gentlemanly, aristocratic assumptions which underlie current practices. One of the students on the evaluation committee strongly questioned the need for secrecy. "The argument," he wrote, "that deliberations must be secret to protect the parties involved speaks little of the individual courage of the people concerned. When will we be honest enough to talk face-to-face?" The wisdom of the ages would probably answer "Never," but youth, the question itself, deserves a lengthier answer. Probably the majority of vexing disputes between faculty members and department chairmen or other officers of administration occur because of failure to clarify what is expected of a faculty member, to set forth the criteria for retention and advancement, or to communicate clearly decisions or judgments of an individual's competence and the basis for such decisions or judgments. Such disputes must take a large share of the blame for the lack of harmony often found within departments whose reason for being is largely that of common purpose and cooperative effort.

Some attempts have been made to counteract these effects by permitting the faculty member more of a chance to make his case: to respond to formally stated written criteria, to discuss his performance with a committee of his peers charged with the review, to be given opportunity for appearance at other stages in the review. Practices of this kind create the danger of excessively formalizing procedures and, at an extreme, of creating an excessively litigious atmosphere. Nevertheless, in an open society, the academic profession can not escape questioning the secrecy in which crucial personnel decisions are reached. My own feeling is that the best direction for reform is to increase the openness and effectiveness of the ordinary means by which an institution takes an interest in and assists in the individual's career development. Part of the fault with existing practices is that they only come into play at the time of reaching crucial decisions rather than as part of ongoing, periodic attention to faculty growth.

Criteria for advancement

A final point has to do with the criteria upon which decisions are based. Throughout higher education, faculty members are judged on the basis of research, teaching, and service, not necessarily in that order. Research at its narrowest means publication. At its broadest, it may be carefully defined to include nonpublished research and scholarship or even to embrace the scholarly cast of mind and capabilities which go into teaching. It is seldom altogether absent in the criteria, even in the junior college, for higher education presupposes advanced study, and advanced study in graduate schools almost always has research as a major part of required work. In the two-year colleges, there is little stress on research and some negative reactions to it. Yet even here, advanced competence in the area of teaching is usually asked for, and the possession of a Ph.D. degree or publication is solid argument for a higher step on the salary scale. I have argued elsewhere of the need to broaden the meaning of research. If the search aspect were more genuinely recognized and accepted, I would not argue with its pervasiveness as a criterion for college and university "professing." As it is, I think it is a loosely defined criterion, often an ill-advised general one, and one upon which judgment is about as loose as upon teaching. Thus in the examination of criteria for teaching, in which student evaluation might have a place, a department might be profitably drawn to examining the criteria for research and trying to improve the means for evaluating it.

Quantitative and qualitative measures

Service is a broadly inclusive term covering most of the tangible services performed inside and outside the university. Administrative service is recognized both within this category and as a separate kind of service. Exact definition is again difficult to set down, though evidences of service appear in almost as tangible forms as evidences of research. A faculty member has either served on departmental or university committees or he has not. His membership in professional societies or in community service organizations is a matter of record. Qualitative judgments are inferable from the importance of positions held in organizations which, like scholarly journals, carry their own qualitative distinctions. Much room for differing judgments enters here, a part of which might be eliminated by trying to make some basic qualitative judgments about the kinds of services performed. Ultimately at the point of specific discussion and decision, attempts have to be made to determine how well a person has performed a service, questions of its general value aside. This is perhaps where the faculty's interest in evaluating administrators might be properly brought forward. Since faculty members make up the bulk of administrators and since so many serve in such assignments part time or for limited periods, refining criteria and judgment within the system of faculty review could go a long way toward a useful system for effective administrative review.

It is commonly charged that teaching is the least well defined and the most inadequately judged of a faculty member's major responsibilities. Were there no truth to these charges, student evaluation would have limited usefulness. But teaching by its very nature is hard to define and difficult to judge. Indeed, the more specific complaint is that both research and service offer tangible measures of accounting — books and articles published; actual services rendered — which teaching does not. Teaching offers only classes taught and these usually assigned by a department chairman or worked out in cooperation with the department. Where few qualitative judgments are made about research and service — that is, pages are merely counted and services recorded — there should be little surprise that teaching fares worse. In such circumstances, all teachers tend to be recommended as "good" teachers. The genuinely superior teacher has little more to show than the mediocre one, and only the genuinely inadequate have to be described in such terms as "good with the superior (or slower or well-prepared or graduate) student." Decisions within a well-defined but not very rigorous system get made with greater weight on the quantitative measures that research and service provide.

On the other hand, as institutions do become more rigorous in their policies, do look at research and service with both clear and clouded eye, teaching does not easily adapt to more rigorous criteria. Publications are read, or if not that, assessed in terms of the editorial rigor of the editors or publishers; research is judged by the selection procedures of sponsoring agencies; service by the nature of who it was for and how important a place the individual had in it. For teaching, comparable evidences might be awards for teaching, the kinds of courses taught, and perhaps special teaching assignments, all of which fail to provide well-defined or widely-applicable criteria.

The point is that whether one examines a somewhat routine system of faculty review or a rigorous one, teaching is likely to be less well-defined and judged than research or service. A department seriously considering the use of student evaluation might see it as a means of giving teaching more visibility and more tangibility within the reward system. Insofar as teaching, more than research and more than service, is what the majority of college professors are principally engaged in, it does not seems particularly startling to insist upon setting high standards for its performance and finding ways of achieving those standards.

Departmental review

Since the department is the place where faculty members have the closest relationships, both with faculty colleagues and students, it seems advisable to think first of the departmental review as the place for student input. Surely there is no great mischief in having the responsible departmental committees formally add student evaluation data to the supporting materials for faculty members being considered for retention, promotion, or tenure. Similarly, department chairmen, if they have not already done so, could add this material as relevant data when considering salaries.

But such an addition would constitute very little improvement in a system which, at best, is orderly, and at worst, casual in its judgments and unfair in applying them. A department's consideration of the place of student evaluation could also be a time for consideration of all its advancement practices and policies and an opertunity to clarify its general educational objectives, policies, and practices. Such a consideration would include the matters mentioned here: review and written restatement of the criteria for promotion and tenure; review of departmental procedures for promotion and tenure; review of on-going practices for developing teachers and improving teaching.

Student involvement

Students can be extremely helpful in these matters of review. The faculty, far from being defensive about involving students, should seek ways of getting their participation, not by one or two or three student representatives serving on committees or being invited to meetings, but by a much larger kind of participation. Few departments are so large that open meetings on key questions cannot be held. Students could be urged to attend, and their ideas and opinions, like those of the faculty, could be listened to and challenged. Preliminary work for such open meetings, hearings, or discussions could also involve students, and follow-up reports could be made by students as well as by faculty. Much of this work might be done on an *ad hoc* basis, the various committees and reports and meetings to end or change functions when the objective — an improved means of recognizing and rewarding teaching — had been reached.

I would hope such involvement would result in other ways of bringing student insights to bear upon teaching and learning. The curriculum could bear scrutiny, as could counseling, as could the relationship between the discipline and the outside world. Probably every department of some size could use an advisory board of students to carry on continuing activities of this kind.

The crucial issue it seems to me, is not the danger of granting the students too much. It is rather in not granting them enough and in neglecting the necessary continuing attention that has to be given if participation in the particulars of teaching, by faculty or students, is going to be forthcoming. There is a good deal more chance that good student evaluations will founder amidst the difficulties of doing a large task well than that evaluations will take over the prerogatives of the faculty.

There are undoubtedly other ways of making use of student input, many of which would be less tied to present practices than these. One can envision a student review board making its own recommendations to a college-wide personnel committee. Such a board might meet with faculty members being considered for advancement. Faculty members might benefit from discussions with students which took place outside of the context of departmental review. Student involvement could certainly prove valuable in making appointments and in clarifying the expectations of the new teacher before he began to teach. The ways are numerous, and they need to be, if only because old ways wear out and new ones serving the old objectives need to be brought into being.

Improvement of teaching and learning

Throughout the discussions of the committee on evaluation, emphasis was placed upon the improvement of teaching rather than upon evaluation as such. The emphasis is proper, for it is easy to lose sight of an objective while preoccupied with a hoped-for means of gaining it. Teaching has too often disappeared in our necessary interest in curriculum, course structure, grades, credit, even methodology. A faculty's interest in teaching is always fragmented by its responsibilities to scholarship, institutional and professional services, university governance, and by the necessity of meeting immediate crises and of reflecting on long-range goals.

The present climate, despite the atmosphere of crisis, seems very favorable to an increased recognition of teaching. The past two decades are marked by a great attention to how students learn and what higher education contributes to a student's development. Teaching can scarcely be talked about these days without having someone object that the focus should be on learning. My own belief is that teaching-learning are two sides of a coin. The whole is what we want, but we can still fruitfully contemplate either side. The project's interest in evaluation forces us to look at both sides, though the area examined is but a small part of the configurations of either teaching or learning. If evaluation can contribute to bringing the campus together in the common teaching-learning enterprise, if it can put some life into daily routines, if it can work specific improvements upon individuals and contribute to the general enhancement of teaching, if it can add to what we know about teaching and lead us to act on that knowledge, then it is surely worth the risks and effort it may involve.

APPENDIX

1

Report of the President's Committee for a Pilot Study in Student Evaluation at Princeton

NOVEMBER 20, 1968

Reprinted by permission of Princeton University

History

In the Autumn of 1965 President Goheen set up a Committee called The Joint Committee of Inquiry on Undergraduate Instruction. This Committee consisted of four members of the Policy Committee and three undergraduates — the President of the Junior Class, the President of the Senior Class and the President of the Student Government Association. This Committee under the Chairmanship of Professor Gillispie spent about six months in making inquiries — consulting faculty members, getting reports about course evaluation efforts in other universities and in some courses at Princeton, considering former Princeton course evaluations, and canvassing student opinion.

In April, 1966 this Committee reported to the President, and its report was printed and distributed to the faculty. The report reviewed the reasons for sampling undergraduate reactions to courses, noted the inadequacies of other surveys at Princeton and elsewhere, and pointed out that these surveys required such systematic and detailed long term efforts and such careful administration to get a sufficiently large sample, that undergraduate administration of such surveys had been uniformly inadequate. The Committee, therefore, recommended that the faculty and administration of the University itself prepare and administer a series of four questionnaires to be given to all undergraduates at Princeton.

The first proposed questionnaire was to be given to students in every undergraduate course. The second was to be filled out by all sophomores to evaluate the experience of their first two years at Princeton. The third, to be given in May of the senior year to each senior, was to assess his reactions to his work in his department of concentration. The final questionnaire recommended by the Committee was to be given to graduates some six to twelve months after graduation and was to inquire about their entire educational experience at Princeton.

In the faculty meeting of the second of May, 1966, Professor Gillispie reported to the faculty and the printed report was discussed at some length. Thereafter President Goheen appointed a second committee to carry out course evaluations in two stages and to report back to the faculty. This second committee was made up of five faculty members with experience in the field of statistics and surveys, two undergraduates (the President of the Senior Class and the President of the Junior Class), an executive secretary, and an amateur chairman. In November, 1966, the Carnegie Corporation made a grant to the University to enable this Committee to carry out its pilot studies. This grant was to pay for printing, for computer time, for released time for a research analyst and for part of the staff of the Office of College Operations, for consultation with the Educational Testing Service, and for overhead.

THE PILOT STUDIES

First stage

The first activities of the Pilot Study Committee involved the preparation of a set of four questionnaires: course evaluation, departmental evaluation, evaluation of the total undergraduate experience, and evaluation of the first two years at Princeton. Different faculty members with experience in such surveys, several Princeton undergraduates, and members of the staff of the Educational Testing Service were consulted at various stages.

When we thought that we had the first three of these questionnaires properly set up with provisions of both rating items and "open-ended" questions for undirected student comments, we arranged to conduct a pilot study at the end of the Spring 1967 semester. We decided to sample courses and seniors in four departments, and we selected Biology, Politics, Electrical Engineering, and French in order to represent the many different content areas and instructional approaches available at the University. We were aware that questionnaires which were adequate for some courses or for some departments might be totally inadequate for others. If we were to be able to make comparisons among the results, however, we realized that it would be necessary to give approximately the same course evaluation questionnaires to students in all courses and to give approximately the same departmental evaluation questionnaires and total undergraduate experience questionnaires to seniors in different departments.

The course evaluation questionnaire was administered at the time of the final examination in 41 undergraduate courses in May, 1967. These courses included all the undergraduate offerings in the Spring in the four selected departments and four more courses in which the instructors had asked to be allowed to try out the questionnaire. The departmental evaluation questionnaire and the evaluation of the total undergraduate experience questionnaire were filled out by the seniors in the four selected departments and by seniors in two other departments which had volunteered to participate in the study. These questionnaires were administered at the time of the senior comprehensives in the Spring term 1967. A sample group filled out questionnaires evaluating their educational experience in their first two years at Princeton in the early autumn of 1967.

Our first objective was to determine whether we could get an adequate response from Princeton students and effective co-operation from faculty and undergraduates. In this respect we were greatly reassured by the first pilot study. 2,169 questionnaires were filled out in the courses in the selected departments, representing 87% of all the students enrolled in the particular courses. The response of the seniors in these departments was also remarkable: we had 220 replies from seniors in six departments, or 90% of all the student seniors involved in these departments.

During the Summer of 1967 the responses to the three questionnaires were tabulated and analyzed. After several discussions the Committee had decided to ask students to sign their names on detachable pages in the course questionnaires so that the Office of College Operations could code the booklets with course grades and with the averages of each student and then tear out the detachable page before forwarding the booklets to the instructors. From the comparisons made possible by this device we hoped to learn to what extent we were getting analytical comments and to what extent general reactions.

The results of our analyses during the Summer of 1967 were as encouraging as the response in May had been. Since the booklets with the students' names attached had first gone to the Office of College Operations and had been held there until all course grades had been turned in, it was possible for us to estimate the association between students' course grades and their opinions of various aspects of the course. We could also consider the students' annual averages in connection with their reactions to courses. These analyses gave us some interesting comparisons.

1. In the more than 2,000 course questionnaires we found only a very slight correlation between a student's evaluation of the course and the grade

2

he eventually received in that course. The correlation between the student's average for the entire year and his evaluation of the course was also very low.

2. When student evaluations of different parts of the questionnaire were compared, we found students making clear discriminations between the parts of a course such as lectures, readings, and precepts. They were not simply blanketing a course with repeated estimates of good or bad or indifferent; they were regularly finding one part of a course good, and another part mediocre, or one part bad and another excellent. Our results in this area were especially gratifying, for one of our goals had been the development of a course evaluation questionnaire which would not simply elicit approval or disapproval (as did so many of the questionnaires we had examined from other universities) but would analyze the degrees of strength or weakness in the various parts of each course.

3. The ratings we received showed considerable variation from course to course; that is, the questionnaire items appeared to be discriminating among the strengths and weaknesses of the various courses. The course evaluation question booklets, with the final pages bearing the student's name torn out were then returned to the instructors involved, together with an analysis of the distribution of ratings for his course, the mean ratings for his course, the mean ratings for courses in his department, and the mean ratings for courses at his number level (e. g. 200, 300, etc.). Individual instructors were also able to read the responses of students to the open-ended questions in their booklets.

Besides the course questionnaires we had three other sample sets to analyze, 1) Evaluation of the four years of Princeton education, 2) Evaluation of the first two years, and 3) Evaluation of departmental work. We can make a few comments on the responses to these additional questionnaires.

1) The four years at Princeton.

a) There is evidence of variation in ratings and in average ratings of students in different departments. The latter suggests that the departments get different types of students or else the departments have different impacts on their students which lead to broad variations in overall ratings.

b) In the section on factors influencing change in the student, senior independent work is listed as by far the most significant experience. While in most cases the factors influencing intellectual growth and those influencing growth as a person are different, the senior independent work is listed high as an influence in both respects. Of perhaps disappointing value, at least in terms of the ratings, are the pass-fail option and contacts with the cultural benefits of New York, Philadelphia, McCarter, and the Theatre Intime.

c) The students do not appear to report an undue amount of negative stress. The absence of coeds seems the most important.

d) In the overall evaluations, the ratings for the academic aspects of Princeton are considerably higher than those for the non-academic aspects. Princeton is rated high on its contribution to readiness for a competitive career, but low on its contribution to a sense of social commitment.

2.) Evaluation of the first two years at Princeton.

These ratings tend to parallel those of the four year evaluation. Again the ratings on academic aspects of Princeton tend to be higher than the ratings on the non-academic aspects.

3.) Departmental evaluations.

a) Again the senior independent work appears as a major positive experience. The senior comprehensives and the junior generals are rated quite low.

b) The advising on senior independent work is rated high while the advising in areas other than junior and senior independent work is rated low.

c) There is clear evidence of variation in the ratings of the six sample departments. Furthermore, since the ratings within a department vary from item to item, one gets a perception of the strengths and weaknesses of a department. For example, for one department the ratings on senior independent work were quite low, but those on the opportunity for useful contact with other students in the department were quite high.

The analysis of student responses to this first round of questionnaires in four (or in some questionnaires six) diverse departments also showed us that certain parts of our questionnaires were ineffective, and we made revisions in order to eliminate duplications or contradictions or questions which turned out to be meaningless for some students.

Second stage THE PILOT STUDIES

The second stage of the pilot study involved a much more complex administrative problem. We wanted to give the questionnaires to all undergraduates in all courses in the University. Not only would the numbers be much larger, but the variety of students and faculty involved would make it more difficult to attain the degree of co-operation which we had managed in the first limited experiment. In order to smooth the way as much as possible we were able to get two stories in The Daily Princetonian to prepare students for the questionnaires, and we sent announcements to prepare the faculty for the administration of the questionnaires.

It was also necessary, because of the large numbers involved in this new survey, to prepare the questionnaires for machine reading. In this we had the co-operation of the Educational Testing Service.

It was with some trepidation that we approached the administration of our course questionnaires to all undergraduates in January, 1968. Our first worry was soon proved groundless, for the response from the students was magnificent. We received a total of 11,910 answer sheets, representing a nearly 95% response from the students in courses participating in the project. 178 of the answer sheets contained invalid course codes and had to be discarded, an error of less than 2% which was kept low by extensive hand checking in the Office of College Operations. Thus in preparing the tabulations and statistics contained in the reports, valid data from 11,721 answer sheets were utilized. These answers were tabulated, and 6,391 report pages were distributed to the instructors, the separate report sheets varying from thirteen to twenty-five pages in length with an average of twenty pages per report.

These reports consisted of a tabulation of the student responses to each questionnaire item. For each item instructors were given: 1) the mean; 2) the percentile of that mean with reference to the average of the responses of students in all courses at the same level (Freshman, Sophomore, Junior, or Senior level courses); and 3) the percentile of that average with reference to the average of the responses of all students in all Autumn term courses. The percentile scores allowed an instructor to determine the ranking of his course on a given rating scale relative to the rank of other groups of insructors on the same scale. Further, for purposes of comparison, averages and distributions of responses (expressed in terms of percentages) were also provided, item by item for 1) all undergraduate courses in the same department, 2) all courses at the same level, and 3) all undergraduate courses in the University. For each of these reference groups two sets of percentage distribution were given: the first was calculated by using all responses of all students in the reference group irrespective of the size of the course involved; the second was computed by averaging the averages of responses by course. Each instructor also received the question booklets in which his students had written comments. These open-ended questions we think important both for the information they give the instructor and for the satisfaction they afford the students in expressing more fully their views of the course.

We also had an opportunity to check the reliability of some of the answers by re-runs. One full class of over 100 students and a number of volunteers in other classes were given their course evaluation questions over again three or four weeks later, after the final examination period. The means for the original testing and for the repeats several weeks later were virtually identical.

We can make certain observations about our experience in these two pilot studies.

1. Our questionnaires need a better coding of departments and course

numbers, perhaps by a system of pre-coding.

2. All the evidence available, especially in the high degree of student response, demonstrates that questionnaires were treated seriously and thoughtfully by the overwhelming majority of the undergraduates.

3. The ratings indicate that, in general, Princeton undergraduates think highly of their professors and their courses.

4. The results show that students can and do discriminate in their judgments of different parts of a course, such as lectures, readings, precepts, laboratories, and exercises.

5. The results also show that the questionnaire items produce satisfactory variability of response over different courses.

6. Reactions from students and faculty indicate that the questionnaire items are meaningful and useful in different kinds of courses and departments; they appear to have excellent 'face validity.'

7. Evidence from an admittedly limited check on the reliability of the responses suggests a very high stability of the average ratings on the different items.

General conclusions

We think there are several general conclusions which we can offer on the basis of our experience in the last two years in developing and administering these two sets of questionnaires, in talking with a fair number of undergraduates, and in listening to the comments and reading the letters of our colleagues on the faculty.

1. Our experience has set at rest our misgivings about student interest in such questionnaires and about the seriousness with which they would be treated. The percentage of response to both sets of questionnaires, and the discrimination shown by our comparisons of various parts of each questionnaire are quite reassuring. During April and May of 1968 many students asked various members of the Committee if the questionnaires would be given again in May 1968 and showed distinct disappointment when they were told that they would not be.

2. Though we have taken no poll, our experience with the comments and letters of the faculty show that those who have reported to us have found the reports of the courses helpful. We think that this helpfulness depends largely upon the elaborateness and detail of our questionnaire. We tried as much as possible to avoid blanket statements about the goodness or badness of the course and to break down the responses to specific features, such as the clarity of presentation of the lectures, their coordination with other parts of the course, the appropriateness of their intellectual level, and so on. Perhaps the Chairman may illustrate by inserting personal experience at this point. I have been through two or three student course evaluations at Prince-

ton in the past, and found the reports either flattering or irritating, but never specifically helpful. In the Autumn of 1968 I was in charge of a large course for about three hundred freshmen. The comments of these three hundred young men I found extremely helpful, and on the basis of their comments I have made modifications in the course. They pointed out to me that one of my lectures in particular was much less effective than the other eleven, and, since I myself had had reservations about this lecture, it will not be delivered again. The students also pointed out that one particular play on the reading list — which happens to be a favorite of mine — was a simple bore to a majority of them. Since this play was clearly failing to accomplish what I intended, it has been taken out of the reading list for the current semester. A significant number of the students pointed out that when I get excited in a lecture I talk too fast. Whether I can remedy this defect or not, I do not know, but these are samples, though perhaps intrusively personal, of what one faculty member has found that student questionnaires can do for him.

Finally, on the basis of many responses from Princeton undergraduates, not only in the questionnaires, but in personal conversations, we are convinced that the administration of questionnaires of the kind we have tried are important in giving students a sense of greater participation in their education. There is no need to belabor the significance of this factor at the present crisis in education. The more we can persuade students to comment thoughtfully and responsibly on their courses and their instructors, and the more we can convince them that their comments are seriously considered, the more Princeton can do for its students and for the cause of higher education in this restless and rebellious age.

Recommendations

These general conclusions and our experience with these devices for student evaluations lead us to make the following recommendations to the President and the faculty.

I. a) That the four questionnaires be given regularly as a part of the Princeton educational process.

b) That the questionnaires be administered for a period of three years, at the end of which period the Advisory Committee and the Office of College Operations report back to the faculty to recommend continuation, modification, or abandonment.

II. That students should not be asked to sign their names to booklets or answer sheets. Our use of names in the two trial runs enabled us to demonstrate the independence of course grades and of averages with which Princeton undergraduates could be expected to evaluate their instruction. Since this independence has been demonstrated, and since a number of instructors have had doubts about the use of signatures and some students may have been inhibited, we think that hereafter signatures should not be requested.

III. That the administration of the questionnaires be made the responsibility of the Office of College Operations or an equivalent office. (This office has done all the actual handling of our preliminary studies; they not only have all our experience, but they have the facilities for handling course registrations, faculty notifications, and the like.)

IV. a) That the machine print-outs include all the information contained in the print-outs of the January 1968 pilot study *except* for the following changes in the print-outs and booklets for the course questionnaires:

1. The percentage ranking of an individual course in relation to all other courses at the same level, and in relation to all other courses in the University should be eliminated.

2. Section 012 (Lecturer compared with all other Princeton lecturers) should be eliminated.

IV. b) That the Office of College Operations distribute the results of the questionnaires as follows:

1. The machine print-outs on individual courses: one copy to the instructor in each course; one copy to the departmental chairman; one copy to a file in a supervised central location for the use of students and other members of the University community.

2. The booklets for individual courses: to the instructor of that course only.

3. The results of the departmental evaluation: the booklets and two copies of the print-out to the departmental chairman; and one copy to the Dean of the College.

4. The results of the questionnaires on underclass education and on undergraduate education: two copies to the Dean of the College; one copy to be made available at a supervised central location for the use of members of the University community.

V. That the President appoint an Advisory Committee of three faculty members and one senior, one junior, and one sophomore from the undergraduate body to consult regularly with the Director of the Office of College Operations on the administration of the questionnaires. This committee would decide the dates for administering, the means for distributing questionnaires and answer sheets, the dissemination of reports, under the recommendations in item IV, and other appropriate matters.

Signed

G. E. Bentley (Chairman)
Robert H. Dicke
Jameson W. Doig
Lawrence A. Pervin

Paul Sittenfeld '69
William A. Stuart
Charles F. Westoff

2

from
COURSE EVALUATION BOOKLET
Princeton University

[The course evaluation questionnaire (which is copyrighted by Princeton University) is in the form of a booklet containing nine parts and a blank page for an essay response. The various parts cover the different kinds of courses at Princeton, and students fill out only those which are applicable. Three of the parts are reprinted here.]

PART 1: LECTURES

1. Applicable ☐ Inapplicable ☐

	Excellent	*Good*		*Fair*		*Poor*	*Not Applicable*
	1	2	3	4	5	6	0
2.-11. Rate the quality of the lectures in terms of the degree to which they:							
2. Held your attention and interest	1	2	3	4	5	6	0
3. Covered the material at an appropriate intellectual level — neither too complicated nor too simple	1	2	3	4	5	6	0
4. Clearly presented the relevant subject matter	1	2	3	4	5	6	0
5. Covered diverse points of view and helped to expand your awareness of alternatives	1	2	3	4	5	6	0
6. Emphasized principles and generalizations	1	2	3	4	5	6	0
7. Stimulated your intellectual curiosity and provoked independent thinking	1	2	3	4	5	6	0
8. Stimulated student discussion outside of class	1	2	3	4	5	6	0
9. Were related to one another and followed a coherent sequence	1	2	3	4	5	6	0
10. Were coordinated with other parts of the course	1	2	3	4	5	6	0
11. Rate the general quality of the							

PART 1: LECTURES (continued)

lectures as a whole	1	2	3	4	5	6
12. In comparison to all other lecturers you have had at Princeton, how would you rate the lecturer in this course?	1	2	3	4	5	6

13. About how many lectures did you miss?
1) 1 or none 3) 5-7
2) 2-4 4) 8 or more

14-22. Often some very concrete characteristics of the lecturer reduce the value of a course or interfere with the achievement of its objectives. For each of the following characteristics indicate whether or not you felt that it applied to the lecturer:

14. Inaudible ☐ ☐
15. Writing illegible ☐ ☐
16. Too abstract ☐ ☐
17. Too superficial ☐ ☐
18. Covers too much material ☐ ☐
19. Speaks too fast ☐ ☐
20. Disorganized presentations ☐ ☐
21. Repetitious ☐ ☐
22. Distracting mannerisms ☐ ☐
(Specify: ....)

Which lectures (or lectures on what subjects) did you find least valuable? Why?
Which lectures (or lectures on what subjects) did you find least valuable? Why?

PART 3: PRECEPTS OR CLASSES

34. Applicable ☐ Inapplicable ☐

	Excellent	*Good*		*Fair*		*Poor*
	1	2	3	4	5	6
35.-44. Rate the general quality of the precepts or classes on each of the following: (The word instructor is used to refer to the person responsible for your section regardless of whether it was a class or a precept.)						
35. Interest of the instructor in the precept (class)	1	2	3	4	5	6
36. Instructor's ability to raise challenging questions	1	2	3	4	5	6
37. Instructor's ability to help clarify readings and lectures	1	2	3	4	5	6
38. Instructor's ability to encourage broad student participation	1	2	3	4	5	6
39. Instructor's ability to conduct discussions	1	2	3	4	5	6
40. Instructor's responsiveness to students' comments and questions	1	2	3	4	5	6
41. Integration with other parts of the course — relevant and supplementary	1	2	3	4	5	6
42. General attitude and preparedness of fellow class members	1	2	3	4	5	6
43. Your own interest, preparation and participation	1	2	3	4	5	6

44. Value of the precepts or classes as a whole to this course 1 2 3 4 5 6

45. How many precepts or classes did you miss?
1) 1 or none
2) 2-3
3) 4-5
4) 6 or more

Were there characteristics of the precepts (classes) which you found particularly valuable or not very useful? Please explain and, if possible, suggest modifications.

Name of Instructor: ....

PART 8: OVERALL RATINGS (Applicable To All Courses)

In relation to your own objectives in this course and compared to other courses you have taken at Princeton, how would you rate this course in terms of how much it contributed to each of the following:

	Excellent	*Good*		*Fair*		*Poor*	*Not Applicable*
91. Your mastery of the relevant content or subject matter	1	2	3	4	5	6	0
92. Your mastery of the relevant skills or methods	1	2	3	4	5	6	0
93. Your ability to see more alternatives and have more insight into the complexity of the relevant subject matter	1	2	3	4	5	6	0
94. Your ability to formulate general principles in the relevant subject matter	1	2	3	4	5	6	0
95. Your abilities for critical evaluation in the relevant subject matter	1	2	3	4	5	6	0
96. An increase in your interest in the field so as to take further related courses or do reading on your own	1	2	3	4	5	6	0
97. An impact upon your emotional sensitivity to the relevant phenomena or upon your values and attitudes towards parts of life or yourself	1	2	3	4	5	6	0
98. Your total educational growth and development	1	2	3	4	5	6	0

Please write in below any generally descriptive comments you would like to make about the course or specific points which were not brought out in the questionnaire.

3

Student Assessment of Teaching Effectiveness

By THOMAS D. F. LANGEN,
Director, Office of Student Ratings
University of Washington

Reprinted from
Improving College and University Teaching, Winter, 1966

Despite doubt of the competency of students to judge good teaching, they *are* the instructor's primary audience. To them he has addressed his communication. If it is unclear to them, he has failed to increase their understanding of course content.

Assessing student opinion of teaching competency involves selection of an appropriate set of items on which judgments are to be made. Even when the selection problem has been solved, some important complexities of classroom interaction are ignored. Many important facets of teaching-learning behavior are peculiar to a particular classroom situation. But the selected set of judgment items permits a comparison, however imperfect it may be, that will satisfy an instructor who wishes to compare himself with a normative group. The set of items may elicit different behavior pattern nuances between a 9:00 o'clock section and an 11:00 o'clock section taught by the same instructor. The difference may be due to real differences between the students in the two sections and may call for a different approach by the instructor. Such differences may be found by two open-ended questions designed to reveal those areas in which the students feel the instructor could improve and those areas in which they feel he performed well.

Assessment of student opinion of teaching effectiveness has been conducted at the University of Washington continuously since 1925. Continued research has resulted in periodic changes of items used. The current survey form consists of a selected set of ten items (see Appendix) and two open-ended questions.

Preliminary to selecting the ten items, forty-one items most frequently mentioned in the extensive literature in the field as characteristics of a good college teacher were analyzed in a year long study to determine the relationship of each item to students' over-all judgment of teaching effectiveness. The ten items which showed the highest common variance with the over-all judgment were then selected. The student is asked to compare his present instructor with others he has had and to rate the instructor on each item as outstanding, superior, competent, only fair, or of less value to him. These value judgments are given numerical scores and the average score on each item as well as on the total of all items is reported to the instructor. For each average score, there is reported a decile standing in comparison with a norm group of the instructor's academic rank.

Faculty participation in the program is voluntary. If an instructor so desires, the results can be forwarded to his chairman or dean but only on the instructor's written request. Over 500 sections are assessed annually and 15,000 students are involved. Students' comments on what the instructor has done well and the ways in which the course might be improved are paraphrased and reported to the instructor.

An instructor may (though few do) elect to have the comments omitted from a report sent to an administrator. Comments, like quotations, may, of course, be altered in meaning when out of context. The context here is the classroom and a bald comment might be misunderstood by an administrator. Since the system of teacher evaluation depends upon voluntary faculty participation and is designed for the improvement of teaching, the administrator's interests take second place. Several have encouraged the use of these surveys whether or not the instructor decides to furnish them with copies. In about one-half of the cases instructors request that the survey be forwarded to an administrator. Some instructors forward all surveys. Others send only their better ones. Administrators like to receive the reports.

A survey customarily takes twenty minutes for a class of thirty students. Surveys are not conducted until the fifth week of a quarter. The heavy demand for surveys occurs in the ten class days preceeding final examinations. Students are advised that results will not be made available to the instructor until after he has filed his grades with the registrar. This procedure protects students from possible punitive action. It enables the instructor to start his new term with the ratings and comments in mind.

Students are very cooperative in answering survey questionnaires. It is seldom that a student will give an instructor the same rating on all ten items. The informal consensus among students seems to be that instructors who request surveys need them less than those who do not request surveys. There is no significant relationship between the rating received by the instructor and the grade the student expects to receive from the course. Since student responses are anonymous, comparison of actual and expected grade is not possible, but class averages of expected and of received grades are quite close. This is true despite the interesting fact that only one student in 5,000 seems to expect an "E" grade.

Open-ended comments and the comparative scores on the items often give direct suggestions on ways to improve or alter teaching traits. There is no direct indication of improvement in understanding or in interest, the two basic teaching purposes. An investigation was conducted during the 1963 summer term* to examine the relationship between the ten items and desired course outcomes of increased understanding and interest. There was a direct statistical relationship (correlation range: .70 to .84) between the ratings on the ten items and the students' perception of the degree of help in understanding received from the instructor. These correlations are higher than those customarily found between high school and college grades, for instance. Correlations between the scores on the ten items and the questions by which the students' increased interest was assessed varied from .54 to .86. In general, trait items which ranked high in improving understanding ranked high in increasing interest in the subject matter. It is safe to say that the instructor who receives the higher decile ratings can have a reasonable assurance that he is accomplishing desirable course outcomes. Additionally, the specific behavioral observations embraced in the item scores and the open-ended comments permit him to alter his classroom behavior in specific ways to increase the yield of understanding and interest.

The investigation was also concerned with the relationship between students' estimation of the instructor's value in helping to understand the subject and such variables as quality of lectures, class discussions, reviews of assignments, and use of syllabi and orientation periods. Using these four elements as predictor variables and the extent to which the instructor helped the student in understanding the subject matter as a criterion, it was found that lectures accounted for nearly 75% of the common variance and class discussion contributed just over 20%. Neither reviews of class work (tests, term papers, and the like) nor the issuance of a syllabus or giving of an orientation talk made any significant contribution. The small contribution of syllabi and orientation talks was surprising in view of considerable comment on how much the students wanted them when they were not given.

This system of teacher evaluation by students has at present many shortcomings. The same items are used for all disciplines and for all levels of instruction. Yet it seems likely, of course, that an engineering drawing instructor and a history professor should be assessed on different teaching traits. In any event, the judgment of the engineering student on the item "Has helped broaden my interests" has a different frame of reference from that of a history student. Yet the decile standing of both teachers, if they are of the same rank, will be based on the same norms. Experienced instructors can assess ratings and comments for their own needs. New instructors may often be confused by apparently conflicting comments. They are free to discuss the reports with the director of student ratings, but he can only address himself to the characteristics of the system. What is needed is a confidant who knows the ratings system and the discipline involved. Unfortunately, such a person is often a senior member of the department and thus a voting member on questions of promotion and tenure. The report with which the new teacher needs the most help is probably the one he would least like the senior colleague to see. The principal shortcoming of the system stems, however, either from the inefficiency of the five point rating scale or the unwillingness of students to give low ratings. As the appended decile tables show, the median associate professor and the median assistant professor are more "superior" than "competent" and the median instructor is "competent plus." Even the bottom 10% of instructors seldom go below "competent minus."

---

*Original research on the ten traits was by Grace French Lazovick in 1957-58.

University of Washington Office of Student Ratings
REPORT ON SURVEY OF STUDENT OPINION OF TEACHING

Name of Instructor —SAMPLE— Course ....
Rank of Instructor Lecturer Comparison Group Instructor
Date Survey Given .... Hour .... Size of Class 42

The ratings assigned by your class have been compared to the norms established from surveys made of the classes of approximately 100 instructors who constitute a representative sample of the 1957-58 faculty of your academic rank. Normative data are available for the following academic ranks: Associate Professor, Assistant Professor, and Instructor. In all other cases comparison is made with the nearest academic rank for which normative data are available.

Listed below is the raw score distribution for your student rating. The decile for that item is shown in parentheses. If you have questions, our extension is 543-1790.

<table>
<tr><th><i>Items on Which Judged</i></th><th><i>Outstanding</i></th><th><i>Superior</i></th><th><i>Competent</i></th><th><i>Only Fair</i></th><th><i>Of Less Value</i></th><th><i>Average Rating</i></th></tr>
<tr><td>1. Interprets abstract ideas and theories clearly.</td><td>| | | | | |</td><td>| | | |x|</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.7380 (5)</td></tr>
<tr><td>2. Gets me interested in his subject.</td><td>| | | | | |</td><td>| | | |x|</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.9524 (3)</td></tr>
<tr><td>3. Has increased my skills in thinking.</td><td>| | | | | |</td><td>| |x| | |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.5714 (8)</td></tr>
<tr><td>4. Has helped broaden my interests.</td><td>| | | | | |</td><td>| |x| | |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.3571 (9)</td></tr>
<tr><td>5. Stresses important material.</td><td>| | | | | |</td><td>| | | |x|</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.7857 (2)</td></tr>
<tr><td>6. Makes good use of examples and illustrations.</td><td>| | | | | |</td><td>| | |x| |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.6485 (2)</td></tr>
</table>

<table>
<tr><th></th><th>Outstanding</th><th>Superior</th><th>Competent</th><th>Only Fair</th><th>Of Less Value</th><th>Average Rating</th></tr>
<tr><td>7. Has motivated me to do my best work.</td><td>| | | |x|</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.7619 (6)</td></tr>
<tr><td>8. Inspires class confidence in his knowledge of subject.</td><td>| | | | | |</td><td>| |x| | |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.4048 (5)</td></tr>
<tr><td>9. Has given me new viewpoints or appreciations.</td><td>| | | | | |</td><td>|x| | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.2381 (9)</td></tr>
<tr><td>10. Is clear and understandable in his explanations.</td><td>| | | | | |</td><td>| | | |x|</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.7619 (3)</td></tr>
<tr><td>On the weighted combination of 10 items.</td><td>| | | | | |</td><td>| | |x| |</td><td>| | | | | |</td><td>| | | | | |</td><td>| | | | | |</td><td>2.6119 (5)</td></tr>
</table>

DECILES TABLES FOR WEIGHTED COMBINATION OF TEN RATED TEACHING TRAITS BY ACADEMIC RANK

Decile	*Associate Professor*	*Assistant Professor*	*Instructor*
10	1.7729	1.9182	1.9575
9	1.9150	2.0125	2.1720
8	2.0230	2.1200	2.2878
7	2.1127	2.1867	2.4800
6	2.1955	2.2800	2.5944
5	2.2782	2.3889	2.6880
4	2.3744	2.4818	2.8020
3	2.5180	2.6333	2.9440
2	2.6450	2.8667	3.1283
1	....	....	....

Note: Outstanding 1.0000, Superior 2.0000, Competent 3.0000, Only fair 4.0000, Of less value to me 5.0000.

Following are abbreviated quotations from the comments made by students in the class. Numbers following indicate how often the comment was made.

SAMPLE

(These are paraphrases of student comments, the number in parentheses indicating the number of students making the comment when more than one made it.)
Note: "Specific Education" is a phrase to disguise the subject to protect this instructor's privacy.

The following comments indicate what your students believe might be done in order to improve the teaching of this course:

Have more class discussion. (7)
A more dynamic approach needed.
Have a clearer outline at the beginning and some guidelines to go by.
Put more inflection in voice. (6)
Vary the course procedure.

Give more continuity to lectures.
Combine outside information in a meaningful manner.
Allow a choice of questions on the examinations.
Assign particular topics to individuals or groups for class programs.
Have more tests.
Use more examples.
Devote more time to the method of organizing and teaching *specific education*.
Show more enthusiasm in the subject matter. (2)
The course title is misleading, for I thought the course would cover more pertinent problems and practical application in the teaching of *specific education*.
Students should know at the outset what is expected of them for the course; that it is the learning of concepts and not history and the overall view that we will be tested on. (4)
Select a standard text.
A history of new *specific education* organization (Peace Corps) might be given.
We should be graded on more than a paper and one test.
The course is dull; make it more interesting.
Follow outline more closely.
Give more direction to the readings.
Voice inflection would improve the enthusiasm of the instructor. (5)
Has to improve his powers of motivating students.
Has to spend more time on present trends in *specific education* and less on history.
Read fewer authors, but more deeply. Enunciate words and speak louder.

The following comments indicate what your students believe you have done especially well in your teaching of this course.

The course was well outlined in the daily presentation. (9)
Lectures were objective and full of good commentary.
Handled a large group of opinionated school teachers very adroitly and kept things moving along. (3)
Has treated us as adults and as individuals.
The material in the lectures was relevant, well researched, and organized.
He aroused interest in the subjects by being interested in it himself. (4)
Brought out some good points in the definition of *specific education*.
Good in his treatment of history and sociology. (3)
Introduced me to a new type of testing which was different and challenging.
He is well-mannered.
Has a mature approach to all phases of the course.
His concern for general concepts rather than minutia increased our interest.
Has given me a clearer concept of the subject matter and has encouraged me to form opinions and arrive at conclusions about it.
Has brought to the class the importance of *specific education* our society and way of life needs. Has consistently contrasted *specific education* in the international scene.
Emphasized the important aspects.
Carefully supports his statements with references and examples.
Created a serene, casual atmosphere.
Extremely good examples were given.
Promoted an interchange of ideas in the class.
Course outline and general instructions are commendable. (5)
He is very unbiased in presenting his views.
Made me think more deeply about *specific education*.

4

University of Washington
Survey of Student Opinion of Teaching

Instructor's Name .... Course and Number ....

My major is .... This course is required .... elective .... (check one)

My cumulative GPA is below 2.5 .... 2.5 to 3.0 .... 3.0 to 3.5 .... above 3.5 ....

This survey is made at the request of your instructor in this class. The information the instructor receives will not identify any student individually. He will receive a summary of class ratings and comments only after the quarter is over. At that time the individual instructor alone determines whether this information is to be destroyed or whether it is to be made available to any other person for reference.

Listed below are several qualities which describe aspects of instructor behavior. Rate your instructor on each of these items by drawing a circle around the number that best indicates his position in comparison with other teachers you have had. Rate each item as thoughtfully and carefully as possible. Do *not* omit items. Of course, it will be the very unusual case when the number you circle is the same for all items.

	Outstanding	*Superior*	*Competent*	*Only Fair*	*Of Less Value*
1. Interprets abstract ideas and theories clearly	1	2	3	4	5
2. Gets me interested in his subject	1	2	3	4	5
3. Has increased my skills in thinking	1	2	3	4	5
4. Has helped broaden my interests	1	2	3	4	5
5. Stresses important material	1	2	3	4	5
6. Makes good use of examples and illustrations	1	2	3	4	5
7. Has motivated me to do my best work	1	2	3	4	5
8. Inspires class confidence in his knowledge of subject	1	2	3	4	5
9. Has given me new viewpoints or appreciations	1	2	3	4	5
10. Is clear and understandable in his explanations	1	2	3	4	5

Your instructor would like to know if there is something you believe he has done especially well in the teaching of this course ....

....

....

....

Your instructor would also like to know what specific things you believe might be done to improve his teaching in this course ....

....

....

....

Thus far your judgments have been restricted to characteristics of the *teacher* himself. For the item below indicate your feeling for the *subject matter* of the course by checking (V) the appropriate entry.

The subject matter or content of the course is: Highly interesting ....
Moderately interesting ....
Not very interesting ....

Circle the final grade you expect to receive in this course: A B C D E

[Note: The above questionnaire was modified for use during the 1970-71 school year by the addition of the following questions:]

11. Lectures gave viewpoints and info text did not contain	1	2	3	4	5
12. Material enthusiastically presented in lectures	1	2	3	4	5
13. Material presented in a well-organized fashion	1	2	3	4	5
14. Helpful to individual students	1	2	3	4	5
15. Integration of material into coherent whole was ..	1	2	3	4	5
16. Text clear in presentation of concepts	1	2	3	4	5
17. Text's overall rating	1	2	3	4	5

18. How much was your interest in the subject changed by this course?
More interested 1 2 3 4 5 Less interested
19. What level of student sophistication was assumed in lectures?
Very high 1 2 3 4 5 Very low
20. Were students free to ask questions, disagree, express their ideas, etc.?
Encouraged 1 2 3 4 5 Discouraged
21. Has improved my problem-solving methods.
Very much 1 2 3 4 5 Not at all
22. Did test questions cover the material emphasized in the text and lectures?
Very well 1 2 3 4 5 Very poorly
23. Would you recommend this course by this instructor to majors in this dept.?
Very highly 1 2 3 4 5 Never
24. Would you recommend this course by this instructor to non-majors?
Very highly 1 2 3 4 5 Never

25.	QUESTIONS 25-27 WILL BE ON THE	1	2	3	4	5
26.	CHALKBOARD IF THE INSTRUCTOR	1	2	3	4	5
27.	WANTS TO USE THEM.	1	2	3	4	5

Students were also asked to give class, age, sex, and expected grade in course.

5

University of Washington
Report of the ad hoc Committee on Student Evaluation of Teaching

Reprinted from Report of Faculty Senate, April 30, 1969

The Executive Committee endorses and recommends adoption of the following report of the *ad hoc* committee on Student Evaluation of Teaching (Professor Donald G. Chapman, Chairman).

Text of the report

This committee was established on January 22, 1968, by Professor Charles Evans, the Chairman of the Faculty Senate to prepare a specific plan for implementation of the recommendation of the Council on Academic Standards for an expanded use of student evaluation of teaching. At its meeting of January 23, 1969, the Senate endorsed both the enlargement of the Committee by the addition of five student members and the enlargement of its mandate to include the following:

(1) A study of what steps can be taken to achieve active and effective participation of students with faculty in the continuing evaluation of the quality and content offerings of the curriculum.

(2) In particular, the preparation of an informational report on participation of students in matters relating to curriculum objectives and content at the department and college level.

(3) Specific policy recommendations on such participation and on the expanded use of student evaluation of teaching.

In response to these the Committee has surveyed the Deans of the Colleges and Chairmen of Departments to obtain the information asked for in mandate (2). We have consulted with Dr. Thomas Langen, Director of the Office of Student Ratings, and had the benefit of his experience with many different rating forms used here and elsewhere in other universities. We have also had the benefit of a rather extensive review provided by some students in the College of Education, one of these by a member of this Committee, Mr. Paul Beisenherz.

Student evaluation of teaching has been going on as long as there has been teaching; carrying out this evaluation in a standard manner and as an official action of the University is a product of the past few decades, and the

University has been in the forefront of this development, due in large part to the leadership of the late Dean E. R. Guthrie. As universities have grown in size it has been necessary to formalize many procedures that had been handled much more casually in the small classical liberal arts colleges. It is only logical that we formalize the opportunity for students to participate in the process of improving teaching, while recognizing that this will be only one form of their participation. Establishing a formal continuing student evaluation scheme will provide a means for feedback from students to instructors and provide one more very visible action in the way of emphasizing the University of Washington's commitment to maximum teaching effectiveness.

Superior teaching is surely one of the goals of the University; another goal is giving any student the widest possible information so that he may intelligently plan his University career with all choices known and available to him. Just as the student needs information from faculty and advisers to intelligently choose between different opportunities, so the faculty needs information and assistance from students on their needs and their desires. Consequently, the University must provide better channels of communication in both directions. As staff, faculty and students work together toward this goal, new and better methods must and will be developed. For example, the concept of the ombudsman, though known for some time in Scandinavia, has only recently been widely known and is still not fully understood or accepted in other countries.

With these basic ideas in mind, to make a commitment to improved communications between students, faculty and administration, we make the following broad recommendations.

I. THAT THE SENATE ENDORSE THE PRINCIPLE OF CONTINUING EVALUATION OF TEACHING: SUCH PERIODIC EVALUATIONS TO BE REPORTED TO DEPARTMENTAL CHAIRMEN AND DEANS, AND TO BE MADE AVAILABLE TO DEPARTMENTAL AND COLLEGE REVIEW COMMITTEES.

II. THAT THE SENATE COMMIT ITSELF TO A TOTAL PROGRAM TOWARDS IMPROVED TEACHING.

III. THAT THE SENATE TAKE STEPS TO IMPROVE CHANNELS OF COMMUNICATION BETWEEN FACULTY AND STUDENTS, AND IN PARTICULAR TO ENDORSE THE PRINCIPLE OF STUDENT PARTICIPATION ON COMMITTEES RELATING TO CURRICULUM OBJECTIVES AND CONTENT, AT BOTH COLLEGE AND DEPARTMENTAL LEVEL.

We also make the following specific recommendations towards implementation of these general policies.

I. a. That the evaluation should be done by the Office of Student Ratings which will have authority to schedule the evaluation sessions in cooperation with the class instructors.

b. That the Office of Student Ratings set up a schedule that insures that every undergraduate class in all colleges and schools (except those now included in a continuing program of evaluation) be evaluated in the University-wide scheme at least once in 3 years and oftener if the budget of the Office of Student Ratings permits. Laboratory or quiz sections that are portions of a class are not included in the planned scheme.

c. The evaluation will consist of the format under development for use by the Office of Student Ratings; a copy of this is attached as Appendix A.

d. To expedite handling and reduce costs it is suggested that responses to open-ended questions be transmitted without editing to the instructor concerned.

The Office of Student Ratings has experience and staff to handle such an evaluation-information program. Alternative approaches were discussed—to turn over the evaluation to an ASUW Committee, to distribute forms at the time of registration or in some other manner. All other schemes appear to have serious flaws and the Committee discarded them from further consideration.

The Committee was informed by Dr. Langen that the frequency recommended (1/3 of all undergraduate sections each year) could be handled within present budgetary limitations. From the viewpoint of individual faculty teaching, say six courses per year, this would mean 2 evaluations per year. Some members of the Committee would prefer more frequent evaluations and this may be possible if the reductions in costs envisaged in Recommendation "d" are achieved.

It was agreed that graduate courses are qualitatively and usually quantitatively different from undergraduate courses. Moreover, instruction in graduate courses is often a different type and with a different goal. For these reasons among others it seems desirable to exclude such courses at this time.

Specific recommendations as part of the improvement of teaching program are:

II. a. Departmental Chairmen should be asked to work with members of their staff on this problem, and in particular evaluations, and the open-ended comments should be reviewed with the aim of seeking improvement of teaching.

b. Departments should be encouraged to set up programs to train and supervise graduate teaching assistants. In particular, steps should be taken to insure greater coordination in those classes where there are both lectures by a faculty member and quiz or laboratory sections supervised by a graduate assistant.

c. Where graduate assistants have full responsibility for classes, it is recommended that faculty members in the department be designated to make visits to all such classes and that such faculty visitors provide written reports to the assistant to help him in improving his teaching ability. Such faculty supervision and reports would be in addition to information obtained through the student evaluation questionnaire.

d. A standing committee of the Senate with student representation should be established to work with the Office of Student Ratings to improve the questionnaire, review special problems of particular departments and possibly work with the Ombudsman in matters pertaining to teaching problems.

Concern has been expressed frequently over the guidance and supervision of teaching assistants; the Committee was informed that the Graduate and Professional Student Senate is studying some aspects of the problem.

While a full study and recommendations are beyond the mandate of the Committee, we do believe that as part of an overall University Commitment to effective teaching, Departments must assume their full responsibilities in this area. The Committee did not attempt to survey all departments that use assistants in their teaching program, but has ascertained that some Departments have instituted useful programs in this phase of their operations. The specific recommendations above are based on such programs.

In its study, the Committee found a surprising lack of objective analyses that demonstrates that student evaluation improves teaching. In part, this stems from lack of research on college teaching and in part from the difficulty of measuring college teaching effectiveness. There are numerous research papers attesting to subjective satisfaction with such evaluations. Thus, the standing Committee set up to supervise the operation might also take as part of its task, the encouragement of research on the whole question.

III. To improve communication and exchange of information between faculty and students it is recommended that:

a. The evaluation form be accompanied by an information form (see Appendix B); information obtained on these responses to this form will be summarized by the Office of Student Ratings and distributed in inexpensive format to Advising Offices throughout the campus. Some copies should also be distributed to Central and Branch Libararies and to some appropriate office in the HUB.

b. Deans should ensure that Departments take steps to prepare advisory information to help students in their choice of courses; such information should be available to other advisers on request and be available to students at the same locations as the course information file.

c. Deans should ensure that Departments prepare examination files and make them available to students. Since most examinations except in very small classes are typed or reproduced by the secretarial staff this could be established by a directive from the Chairman to the secretarial staff at trivial cost and additional effort.

d. The Committee believes that departments should review their curriculum at periodic intervals and that students at both graduate and undergraduate levels should be invited to participate in such curriculum reviews. Student participation may be more useful and meaningful on such ad hoc committees than on standing committees which tend to be involved in routine matters.

The Committee discussed at some length the Course Catalog inaugurated in the College of Arts and Science in Autumn, 1968. It is not clear at this time how useful or how adequate this catalog is. The Committee felt that difficulties might arise in obtaining responses from individual faculty on a continuing basis and believes that more useful and more up to date information can be obtained through the simple addition to the course evaluation form of the course information form and by information pamphlets from Advisory Offices (as recommended in II, b). We recommend placing responsibility on the Advisory Offices since advisers are familiar with the need and will have a self-interest in the information file from both sources (i.e., from students via the Office of Student Ratings and from other advisers). A number of information pamphlets now being issued by Departments were reviewed. Some of these are limited to discussion or course changes; others are more complete and valuable for advisers.

Appendix C is the informational report on participation of students in matters relating to curriculum objectives and content. Information at the College level was obtained in part from Vice Provost Eugene Elliott, and also through questionnaires; at the Departmental level through questionnaires; except in the College of Arts and Science. Professor Morton Kroll conducted a survey of Chairmen in Arts and Science during winter quarter and since one part of his survey gave some information on this matter, it seemed undesirable to burden respondents with additional survey material. The committee expresses its thanks to Professors Elliott and Kroll for making available the information they obtained.

While few Departments as yet have made formal arrangements for student participation on committees, many Departments have some informal arrangements and, of course, quite small Departments can handle this on an even more casual basis. Summarization of some of the informal mechanisms has been difficult, and the Committee realizes that some misrepresentation may have been introduced inadvertently in the process of summarization. It is also important to point out that this is an area that is changing rapidly—our report shows the situation as it was in winter quarter 1969.

Douglas G. Chapman, *Chairman*
Mr. Paul Beisenherz
Professor Allen L. Edwards
Mr. Kjetil Flatin
Mr. Richard Heller
Professor Robert Hennes
Mr. Peter Klein
Miss Melinda Lockett
Professor Scott H. Lytle
Vice Provost Marion Marts

April 30, 1969

6

How Visitation Came to Carnegie-Mellon University

by Arthur M. Eastman, Carnegie-Mellon University

(Reprinted from *Bulletin of the Association of Departments of English*—May 1969)

When the new chairman of the English Department came to Carnegie-Mellon from the University of Michigan in the fall of 1968, he came to a department that did not visit its teachers from a department that did. The system at Michigan, practiced during the first four semesters of a new man's appointment there, had its drawbacks, of course—bland reports from lazy visitors at times, hyperactive adrenalin secretion among the insecure, occasional obtusenesses on the part of visitors or conflicts of personality between teacher and visitor—but by and large it was useful. It improved pedagogy and it gave the department some more precise sense than it otherwise would have had of the abilities of its men. It was a way of practicing institutionally the ideal of the examined life. So the new chairman conceived it, at any rate, and he wanted to see his colleagues adopt such a system at Carnegie.

His elder colleagues soberly saw some difficulties. Not only would there be the usual human disinclination to submit to evaluation and the usual fear that evaluation would be myopic, prejudiced, and ruthless, but to institute the scheme properly probably meant visiting all the instructors (12) and all the assistant professors (10)—22 out of a department of 38—a taxing job. Though they did not fully communicate the point, they saw, too, that if visitation arrayed or seemed to array the younger as the visited, their elders as visitors, it would activate and intensify schismatic tensions that as recently as the preceding spring had threatened to disintegrate the department.

Notwithstanding, the chairman thought the department could be persuaded to accept the practice. Nine months before coming to Carnegie, at the interviewing that led to his appointment, he had indicated that he regarded visitation beneficial. In the fall he had brought up the topic at various meetings with members of the junior ranks. To members of the department he circulated copies of the forms used at Michigan, one notifying the instructor of the impending visitation, the other notifying the visitor of his duties. These forms, he indicated, had to his mind about the right content and tone:

> Dear Mr..:
>
> I should like to ask you to visit this semester classes taught by Mr.. He will be asked to arrange dates with you and to acquaint you with his handling of course materials.
>
> The principal object of these visits is to give encouragement and friendly counsel to new or relatively new members of our staff. To this end, it is indispensable to confer with the instructor you are visiting and to give him appreciative and candid comments that may increase his effectiveness and his sense of professional achievement. Although our visiting program emphasizes classroom performance, the visitor should also take into account less obvious but important matters such as the instructor's ability to interest his students in the basic objectives of the course, his handling of written assignments, and his professional competence generally.
>
> After your final conference with the instructor whom you have visited, will you please write for the Department a summary of your impressions.
>
> Yours truly,
> Committee on Lower Division Courses

> Dear..:
>
> It is the policy of the department to ask senior members to visit classes of a new member during each of his first four semesters of teaching at Michigan. The principal object of these visits is to offer friendly and professional counsel; another object is to provide in a memorandum from each visitor to the department, a record of informed and responsible opinion, representing several points of view, for the use of the Executive Committee in its discussions of reappointments and promotions.
>
> The interests of the visitor go beyond classroom procedures. He will be glad to discuss many aspects of the course: planning, paper assignments and marking, examinations, the place of the course in the curriculum. Normally, he will confer with you after visiting two or three classes early in the semester, then visit again later on.

This semester I am asking... to be your visitor. Please confer with him soon to arrange the dates of his visits and to acquaint him with your handling of course materials.

Sincerely yours,
Chairman, Committee on Lower
Division Courses

Finally, he brought his proposition to the departmental Steering Committee. This body, consisting of three from the junior ranks, three from the senior, had come into being early in the fall at the instigation of the chairman, who had wanted to be chairman, not head, and to govern his department not from above but by the side of, working with his fellows toward common goals. He had seen a democratically elected Steering Committee as a way of institutionalizing this relationship. At first the department had faintly resisted, legitimately Missourian about its value, but it had finally gone along. And the promise of quicker and broader dissemination of information and of more general participation in departmental government had seemed to be in the process of being realized. The Steering Committee was, therefore, the appropriate group to take up and settle the question of visitation—unless the entire department was to be convened as a committee of the whole. Past experiences at Michigan and Carnegie dismissed this alternative from most minds.

Discussion was vigorous. The instructor on the committee, knowing that the department was planning to contract its Freshman English offerings and hence its staff and that no instructor could hope to stay on beyond his allotted three years, saw that visitor's reports might well strengthen and particularize the all-too-commonly bland terminal letters of recommendation, and he spoke out firmly in support of the proposal—firmly but briefly. The others were, to a man, not hostile but doubtful, and the chairman, challenged, allowed himself to argue passionately for his beliefs. Why assume the visitor would be rigid, fat-headed, impervious to the new? Did these adjectives apply to the members of this Committee? Why not perceive the visitor in the image of themselves? If they were sensitive, ought they not attribute sensitivity to the visitor? Endow him with their own generosity, kindness, openness. See him in the mirror of themselves, their best selves.

The upshot of the meeting was threefold. Most obviously, there was the appointment of a subcommittee to work with the chairman upon the question, to see whether certain doubts or alternatives might not be effectively dealt with. Less obviously to the chairman, most members of the committee still harbored very grave doubts and did not, in fact, envision visitation as actually coming about. Less obviously to the committee, the chairman, persuaded of the success of his rhetoric and confident of the merits of his proposal—the chairman did envision visitation as all but present, approved, in action. The lone instructor may have seen all this. He seemed to be enjoying himself.

These weeks brought to the chairman's desk a quartet of thoughtful memoranda, most of which had also a wider circulation within the department. They expressed concern lest visitation be authoritarian, bureaucratic, insensitive, inhumane, stultifying, corrupting, divisive. They proposed alternatives: an "open" system with all classrooms subject to visitation by any member of the staff, pedagogical discussion groups meeting regularly or irregularly, team teaching assignments, the separation of evaluative from instructional operations. The arguments can be extrapolated fairly easily, but it may be useful here to let the memoranda speak for themselves. Each expresses not so much an idea as a person whose various private, professional, and public concerns all came complexly and thoughtfully to bear on the question.

The first was by Assistant Professor Roger Burbridge, (aged 35, Ph.D. Connecticut 1967, entering his second year at Carnegie-Mellon):

> Visitation as you conceive it (Michigan style) runs against the healthy liberalizing trend in American education, and threatens to institute a sterile authoritarianism at a moment when the department can ill afford it. The system is envisioned as promoting mutual respect and cooperation among the ranks, but in my opinion it will destroy the good will that exists; it is supposed to provide younger teachers with the helpful experience of their seniors, but I fear that it will produce humiliation and resentment. Young teachers need no reminder of their inferior and precarious status, especially from a system which could so easily become punitive rather than instructive.
>
> The idea that, in a rapidly growing society, wisdom is peculiarly invested in age is logically indefensible and even morally repugnant. The growing gap that exists between our students' needs and what, traditionally, we have given them can only be bridged by new ideas, experimentation, even by a complete rethinking of what education should do; such innovation springs typically from young teachers, and we should not stifle their creativeness and enthusiasm—an inevitable outcome, I believe, of a Michigan-type system. Even a teacher's errors can be valuable; a teacher who does not make mistakes and learn from them must be singularly tradition-bound and unimaginative.

> I urge you to consider a plan that would incur the minimum of rancor, would be genuinely educational, and would realize all of your constructive goals as well: an open-doors policy which would give all teachers freedom to visit one another at their mutual convenience. I can gain more from visiting a superior teacher, be he full professor or instructor, than from the unsolicited judgment of a senior staff member with whom I may have profound pedagogical disagreements or who even—and it happens—be personally hostile to me. And a visitor might even learn something from me—a possibility he could overlook if he were sitting in judgment. Lastly, your purpose of evaluating teachers for promotion would be served far better by soliciting a large number of informal views from those of their colleagues who have seen them in action.

The second was by Assistant Professor David Demarest (aged 37, Ph.D. Wisconsin 1963, entering his fifth year at Carnegie-Mellon):

> I am opposed to the Michigan plan of class visitation—without careful re-definition of its applicability here.
>
> My rationale is like this:
>
> There are two primary purposes for class visits—
>
> 1) Evaluation of teachers who are up for career decisions;
> 2) Exchange of pedagogical wisdom.
>
> If the first is our prime objective, let's not set up an elaborate bureaucratic system. Let's just simply each year appoint three or four people who will have a responsibility to visit and evaluate.
>
> If the second is our objective, we can do much better than Michigan.
>
> Let's recall first that teachers learn from watching teachers. That's the great method of learning what it's about—helpful advice by visitors to the contrary. Any visiting system should thus involve reciprocity between faculty members.
>
> Second, if teaching wisdom is the end, reciprocal visits among peers will be as important as those between ranks. If we must institutionalize a system, let's recognize that fact.
>
> Third, we do at Tech have audio-visual replay taping available as a method. Let senior and good teachers tape a couple of classes a year and invite junior members to a showing. Let younger staff reciprocate with a tape of their choice. Talking over the comparison will be helpful to both staff members.
>
> But rather than any system, why don't we declare an "open" department—any staff member, any student welcome to any class, always. Learning after all is public, and professionally we're com-

> mitted to the sharing of our views. Let's publicize to the students that our classes are always open. Let's invite faculty, of any rank or department, if they are interested.
>
> Let's reject red tape—or evaluative motives hiding behind interest in pedagogy. Let's open it up.

The third was by Assistant Professor James Hastie (aged 33, M.A. New York University 1961, ABD Indiana 1966, entering his fourth year at Carnegie-Mellon):

> Something there is about the tone and quality of the "Michigan letters" which bothers me. At least, they seem stuffy; moreover, I feel they are vaguely offensive. The ends of improved instruction and faculty evaluation might be better served in other ways.
>
> First of all, I regard education as an ongoing experiment. I have never taught the same course twice. I always make mistakes, and sometimes I recognize them and learn from them. I discuss my teaching, my classes, and my materials with my colleagues often and at length. I have no objection to permitting fellow faculty members to drop in and watch my classes. Indeed, as Dave Demarest put it about his 12th grade class, "I need all the help I can get!"
>
> What bothers me is the exclusiveness of the evaluation procedure. In some of my classes I use a frankly open and experimental approach. I can well imagine that some of my more traditionally oriented colleagues might be startled by the experiments. In my own mind I can already hear the two most likely responses: either madness or an abdiction of responsibility. While both of these responses probably have a basis in fact, the more important fact is that I encourage my students to be critical of me, of each other, and of themselves.
>
> Again, I have no objection to visitations. What I object to is the nature of evaluation. If a traditionally oriented colleague is going to visit me in class, my impulses would be to "perform" in the classroom whatever it is that teachers are "supposed to do." Such pressures toward "performances" strike me as intellectually dishonest and pedagogically unfortunate.
>
> Perhaps a better way to work for more effective teaching would be for the department to have a weekly happy hour where we could get together specifically to discuss what is going on in our classes. A loosened up, free discussion (a lush's guide to better teaching?) of content, methods, responses, etc. seems more likely to result in understanding and communication than an occasional visitation and chat.

> The end of evaluation might be better served by "open" classes, essentially along the lines suggested by both Dave and Roger in their recent memos. I am willing and even eager to share my "happenings" with curious or interested faculty—and even students! But I would also like to think that evaluation of my laboratory could also reflect a wide spectrum of reaction.

These memoranda were from the junior ranks, from those who might regard themselves as the victims of the visitation. But the fourth came from Associate Professor John Hart (aged 51, Ph.D. Yale University 1942, entering his twenty-third year at Carnegie-Mellon) whose comments, the chairman was later to learn, seemed to many in the upper ranks to settle the matter.

> I suppose I ought to put down on paper my feelings about visiting classes while a statement will still be of use.
>
> The trouble with firm resistance to visitation of classes is that it smacks of fogeyism (since we have never done it) and it denies the obvious claim of quantitative measurement without which anyone automatically and permanently classifies himself as dinosauric.
>
> Let me start by saying that I am not trying to keep the door of my classroom shut. It has always been open—to rare auditors, colleagues, visiting firemen, stray dogs and an occasional pigeon. The visits, I am sure, have never been instructive—on either side. The visitor has been treated to a slow-paced, rambling, non-directed discussion of what he feels to be nothing terribly consequental. I have the advantage of mumbled thank you's followed by a hurried disappearance of the visitor (presumably to the nearest men's room to throw up). That's all right; the door is still open, preferably *without* advance notice, to everyone from the janitor to Stever [Carnegie-Mellon's president], as long as he doesn't keep stamping his feet.
>
> The objection is more generalized. It is out of concern for the members of the department, out of consideration of what the relationships between those members have been and must be that I am against the whole idea.
>
> Let's consider the purported advantages of visiting:
>
> 1. It helps guide and improve the young man's teaching. Visitations are selective and occasional; that means that they are special and tense hours for the instructor and that they are fragments of the course and probably scattered hours for the visitor. Such visitations bear, I maintain, little resemblance to the course as a course taught by the instructor over a semester or even to his day by day performance in the classroom. The "guidance" is *likely* to be limited to

mannerisms and superficial techniques rather than anything essential to the relationship of teacher and students; it is further inevitably influenced by the visitor's *own* methods and habits which no doubt suit that person very well but would be completely incomprehensible or wrong for the instructor. Anything gained in the way of improvement would seem to me to be negligible and worth neither the time nor the anguish.

The only way an instructor can learn is by going in and trying and failing and going in and trying something else and failing and so on, until he himself has over the years established a rapport with his students based upon his own personality, his knowledge of his subject matter, his developed ability to listen to his students, to come to know when to push ahead and when to pull up short, etc. Even then, even with the best teachers in the world, classes can go wrong, too great a gap may occur.

2. It provides evidence for evaluation of the young man's teaching. I don't mean to make ogres of visitors here, or to insist on the spectre of fear at using visitation for evaluating purposes; the instructor inevitably invites his own fears, develops his own lump in the throat at the prospect. What I do today in class may determine what my visitor says about me, what my future recommendations are, what place I am able to teach at, my promotion, my whole future. The man who is not able to conjure up such images is not worth his salt; the man who masters all these terrors has mastered all these terrors—he has not indicated anything about his ability to teach. He might get an A from Dale Carnegie, though, and from his visitor, too. In short, the evidence is there but for reasons suggested in item 1 and added to here, it is quite likely to be highly unreliable. It looks good, it is something on paper and because it is on paper it is going to impress administrators. But let's impress them with the right things.

If we reject visitation, what do we have to go on?

1. We have had joint teaching experiences before and will have again. I have taught courses with Sochatoff, Slack, Woodruff, Cottrell, and Hayes. I also taught a semester course jointly with Dave Eskey in his first try as a teacher; Slack did the same with Bob Crozier.

2. Encourage semester-long course-taking by colleagues (not for credit). In my time, I have sat in on sections of courses taught by Parshall, Schutte, Steinberg, Goodfellow and Alemany that I can think of offhand. It took some extra time, but it was time well spent.

> 3. Arrange discussions of novels taught in the larger courses at which instructors can consider meaning, approaches, questions, whatever. And discussions lead to all kinds of ideas and implications for people engaged in them. They have always been helpful.
>
> 4. Throw people together a lot. English teachers are going to talk shop if given any chance at all. Free exchange leads to experiment and improvement.
>
> All the above may and I think does contribute to the stimulation and improvement of teaching. It also gives full opportunity for including evaluative judgments of teaching. Over a three year stretch (which is the usual time we have instructors), any kind of contact with a man will reveal whether he is working at, thinking out, ideas. It is far more dependable than the coffee room gossip Jim Myers was predicting as a basis for evaluation. It is in my judgment sounder than visiting, however contradictory that may seem. And it allows for a more natural kind of relationship (friendship, even) than the supervisor-employee thing that visiting it seems to me necessitates.
>
> With 8,000,000 instructors (about right for Michigan?) visitation is the only way. In a department with 35 people, five of whom are new, no.

It began to be borne in on the chairman that to his new colleagues the University of Michigan loomed as an educational River Rouge plant, gigantic, industrial, and impersonal. But though he sympathized with their doubts, he thought he detected in certain paragraphs of certain of the memoranda a tendency to arrogate virtues to the self and impute vices to the system. And the argument variously urged for an "open" department he viewed doubtfully. If it meant all classes to be open to visiting at all times, it seemed to him a far graver intrusion on privacy than was the system of visitation he proposed. If it meant something less extreme, then it was a fine idea but to be viewed rather as a supplement to than a substitute for visitation. So it was that he met with his subcommittee members, argued anew according to his old persuasions, concurred in certain modifications of the Michigan arrangement (of which more hereafter), and achieved, so he thought, further if somewhat reluctant approbation for his cause. The next day department members found in their mail boxes, over his initials, the following notice:

> Shortly the department will inaugurate its visitation program. The plan is to visit all instructors both this semester and next, all assistant professors either this semester or next.
>
> I know there are doubts and some dismay. I know that visitors

> have to spend time they might use elsewhere. I know that the person being visited may suffer anxiety and harbor resentment. My conviction is that the gains are far greater than the losses in improving teaching, in providing ourselves with that kind of precise knowledge about our own performances which will help us in matters of promotion and placement.
>
> There will be questions, of course, which I shall be happy to talk over at all times.

According to its regular schedule, the Steering Committee met that afternoon. Its members had one question, in effect, though it took many forms: Did the chairman think that the committee, the week before, had indicated its full approval? Was not the subcommittee to report to the parent committee for approval or disapproval? Was this newly instituted Steering Committee to steer or merely rubber stamp? The emotions inspiring the questioning were varied, of course, but intense. The dominant ones were anger and dismay—not so much about the visitation program, which became at this point a subordinate issue, as about the whole nature of the departmental community and its government. Was it to be authoritarian or democratic? The chairman's own emotion was dismay at the failures of his judgment and sympathy. The approbation he had ascertained was deference, was courtesy —even compassion, perhaps—not approval of his position. His sense that everyone had talked, argued, and had been reasoned into consent and that more than sufficient time had been allowed for the process derived from an illegitimate projection of his own feeling for the case.

The committee and its chairman were having a confrontation but no one wanted victory. Each wanted a viable solution, a way to resolve this issue and promote the department's working democratically together. Wanting this, they worked toward it indirectly arguing again fervently the causes for and against visitation. The respect and decency with which they conducted the argument were among the causes for its ultimate resolution. Another was the pair of modifications of the Michigan program that the subcommittee had adopted. One of these was to require that the teacher being visited receive a copy of the visitor's report. The other was to let him ask for further visiting if he felt the initial visitor to have been negligent, prejudiced or otherwise unjust. As the committee reviewed these matters, it could see that under the Carnegie-Mellon program, the person being visited was not quite the hapless kind Michigan had been seen as reducing him to. Yet, no agreement, and the usual time for an adjournment was in sight. It was unthinkable that the issue go unresolved to canker the spirits of all departmental members that night and thereafter until peace was lost from among them. At that point the chairman offered this choice: to go forward with the visitation program, committed to it, or to veto the entire operation. Decision would be better than no decision, and if it were negative, well, chairmen have to learn how to live with embarrassment. The committee members agreed, but they rejected the chairman's offer to leave the room while they conferred. He must stay. Then, one by one, they chose, each with his own reservations—to go forward with the program.

It was a generous resolution. Warmed by it and tender still from the emotions that had gone into its shaping, the committee and the chairman stepped into the outer office to extract a last minute notice from their mailboxes and go home. What each found in his box was the following memorandum:

TO: Members of the English Department
FROM: F. Dennis Williams (aged 24, B.A., Wisconsin, 1965, A.B.D. Rice 1968, entering his first year at Carnegie-Mellon)
RE: VISITATION

"I have lived some thirty years on this planet, and I have yet to hear the first syllable of valuable or even earnest advice from my seniors."

—*Walden*

While all of Thoreau's readers over the age of thirty may be inclined to disregard his opinion, there is little question that it is a belief not uncommon among young and rebellious staff members, particularly in the more enlightened American universities. It is, in part, the reason for my objection to a procedure which appears to have been accepted in great haste by this faculty. Visitation has no place in the revolutionary educational processes of this decade; it shall return this institution to the limbo from which it has so recently struggled to emerge.

The nineteenth-century educational system in this country was based on a very strict form of authoritarianism. In the public schools, a movement to supplant dictatorial supervision began as early as 1900, and became a prominent movement after 1910. It was then the claim that supervisors, rather than providing teachers with a variety of suggestions and a wealth of materials from which they might select, had become a kind of inspector, laying down prescriptions regarding methods and materials. With the decline in power of supervisors, the system of formal evaluation by teachers and administrators gained support. It is only in the past thirty years that this system has receded. In the secondary schools, where the experience and education of the teacher are more readily open to

> question than are those of the professor, formal teacher evaluation is no longer the accepted practice.
>
> In the university, the system of visitation is nearly unknown. The anachronistic existence of this institution at the University of Michigan seems hardly sufficient precedent for its precipitate acceptance by our department. As early as 1915, the National Education Association adopted a resolution in opposition to "those 'ratings' and records which unnecessarily disturb the teacher's peace and make the rendering of the best service impossible." In recent years, the NEA has restated its objections to all plans which require "subjective ratings." Is there anyone on this faculty so naive as to believe that this proposed system will be objective? With the knowledge that the foremost educators in this nation have rejected such forms, and in addition have singled out unstructured comments as a source of subjectivism, shall we nevertheless institute this program of visitation within *two weeks* following the opening meeting of the steering committee. It is a characteristic of wisdom not to do desperate things. Where is that reasoned discussion supposed to characterize the academic world? "Wisely and slow; they stumble that run fast." (*Romeo and Juliet,* II, iii, 94)
>
> It is apparent that this system is ready to begin operations. Memoranda written by younger members of this department have been treated with levity; thoughtful contemplation of the more serious issues has not taken place, and, one must assume, will not take place. I encourage all members of this faculty to refuse to implement this system until it has received a full and adequate hearing. I refuse to participate in such a program, not only because of the haste in which it has been conceived, but because it is a violation of every principle of academic freedom and will inevitably lead to standardization and the limitation of that experimentalism so vital to a developing university. Any man more right than his neighbors constitutes a majority of one.

That night and the next day knots of concerned department members collected in the corridors and offices of Baker Hall. Should there be revolt? Was Big Brother here to stay? There was that kind of thing and much, much more that was less fearful, more constructive. Should there be further discussion, a departmental meeting, etc.? Several members came to the chairman to review the entire matter. Through the hours of the day the uncertainties dragged on, the chairman and the Steering Committee's voice clear, the department's irresolute reluctance equally clear. Then again in the mailboxes appeared a memorandum, this from another instructor (James Myers, aged 24, M.A. Columbia 1966, entering his second year at Carnegie-Mellon), the instructorial member of the Steering Committee.

MYERS *contra* WILLIAMS

> "I mean what would you rather do. What comes into your head first. No matter how silly it is."
>
> "I don't know," Cohn said. "I think I'd rather play football again with what I know about handling myself, now."
>
> —Ernest Hemingway

I have known people who thought Robert Cohn was the hero of *The Sun Also Rises,* but I have always considered them fools, even though we were all young staff members teaching the novel to freshman at the enlightened university of Cleveland State. They hated Hemingway for his cynical defeatism and were much happier teaching *Brave New World,* where the enemy, the establishment, is less elusive; and the hero, Savage—a great admirer of *Romeo and Juliet*—more heroic in his determination to do his thing. These indeed were young revolutionaries who, unlike the old revolutionaries, were astounded at the prospect of anyone disagreeing with them, much less standing them against a wall.

Nothing short of an apocalypse would have persuaded them that they were responsible for what they did in class and that they should ever have to answer for it. "A Last Judgment," Blake said, "is Necessary because Fools flourish." I have always been a great admirer of Blake.

But enough of these rambling, senile remembrances of the academic year 1966-67; let me discuss the problem at hand: visitation at Carnegie-Mellon. A fearsome word, visitation, with its supernatural connotations. It raises the spectre of force, of the instructor's being forced to become a plastic model of some fusty old fart of a professor. How does that happen, I wonder? *Because* the old man threatens to give us bad reviews if we fail to conform? I cannot see the causality, the inevitable necessity. I cannot see howling for the cloistered, inviolable sanctuary of the instructor's classroom; "that which purifies us is trial, and trial is by what is contrary." The instructor who would conform to escape censure is venal, callow, cowardly, regardless of the power marshalled against him.

"Hige sceal þe heardra, heorte þe cenre,
mod sceal þe mare, þe ure mægen lytlað."

("Purpose shall be the firmer, heart the keener, courage shall be the more, as our might lessens." Donaldson, trans.)

But perhaps no one is impressed by the Anglo-Saxon ideal; perhaps we have lost the heart as well as the language. What cheer for the coward, then? Let us consider what tangible things we could lose through unfavorable evaluations. Tenure? Promotion? As you will recall, instructors of English may stay at Carnegie-Mellon for three years. At the end of three years, barring divine intervention, we leave. Further, as Mr. Eastman has observed, the department needs no justification for refusing to extend anyone's contract. "When you ain't got nothin, you got nothin to lose." (Bob Dylan, "Like a Rolling Stone.")

That we stand to gain by favorable evaluation seems to me tediously obvious—and unfavorable reports will not remain in our dossiers when we leave C-M U. The possible harm is negligible and the possible gain finite, a consideration that ought to inspire some tepid belief in the fearful: "cela vous fera croire et vous abetira."

I have my doubts on the efficacy of visitation for improving teaching; I am of John Hart's opinion that we learn by falling on our faces. I am not going to change my methods unless I believe down in my gut that I have been mistaken. But I am willing to listen to advice, even from my seniors; and if I were not, I should have some difficulty justifying my choice of profession.

The academic world is not predicated upon the assumption that every man within it is a phoenix, and academic freedom is not the freedom to be a law unto oneself. The world and that freedom will inevitably lead to the worst standardization of all: every man is right. That notion goes beyond the childish Haroldian pose of rebellion, beyond the case of arrested development, beyond the pale and into the abyss where we no longer have the freedom to assert that "a line is a line in its minutest subdivisions," to assert that a thing is itself and nothing else. When we are all right, then we are truly "all inter-measurable by one another: certainly a happy state of agreement, in which I for one do not agree. God keep you and me from divinity of yes and no too—the yea, nay, creeping Jesus—from supposing up and down to be the same thing, as all experimentalists must suppose."

To hell with your revolution, Williams.

After such passion, what delay? Lists were drawn up, names assigned, and visitation in the English Department at Carnegie-Mellon University got under way the following week. Whether it would work, or how, or how well—these were questions to be resolved hereafter.

Sequel, Autumn 1970

Developments to data: 1) During the remainder of 1968-69, visiting took place and, the chairman believes, usefully. Visitors were generous yet on their mettle; one or two of those visited asked for reviews by other visitors or for opportunities to speak or write rebuttals. Most were encouraged at the approval they received or at the quality of performance they witnessed. 2) During 1969-70 the Department established an evaluation committee with representatives from each rank and empowered it to offer to *all* staff members, senior as well as junior, a choice of ways of being evaluated: one might evaluate his own course, be evaluated by a visiting team, or have students do the job. Or one might pick two or three of these methods. Most opted for the student evaluation, and most here settled for the A Phi O form in general use on campus. Two wrote extraordinarily detailed, sensitive, and exemplary self-evaluations. And some three or four chose to be visited and were, reports being written by several members of the visiting team.

3) This year we shall review our operation and try to improve upon it. The chairman is eager to help new staff members and conceives visiting as one way of doing it. Increasingly, though, he sees the need to keep the judgmental aspect of visiting down, the instructional or coaching up. He finds, too, that many of the older young staff will accept visits from him as a function of his position while resisting visits from others. He intends to visit quite a lot — where welcomed. The chairman of the evaluation committee, meanwhile, trusts to develop a far more sensitive evaluative instrument for student use, relying on last year's PICT conference on evaluation and on inquiries and researches conducted since. Even where none of these efforts directly produces that improvement one hopes for, there may be a Hawthorne effect.

7

From Effective University Teaching And Its Evaluation

Report to Academic Senate of the Davis Campus of the University of California, April 1970

Milton Hildebrand and Robert C. Wilson, Coordinators

SUMMARY

Our objective is to contribute to the improvement of teaching at the university by characterizing effective performance and providing a satisfactory basis for the evaluation of teaching.

Four surveys were conducted: 338 students described the teaching of teachers they regarded as their best and worst instructors; 119 members of the faculty described the teaching of colleagues they regarded as best and worst instructors; 162 members of the faculty described the ways they distributed their time among various academic pursuits; and 1015 students independently described the teaching of instructors previously rated by other students and/or faculty as to effectiveness of teaching.

The principal results follow. [To conserve space here, the supporting documentation, discussion, and data given in the full report has not been included.]

1. There is excellent agreement among students, and between faculty and students, about the effectiveness of given teachers.

2. Best and worst teachers engage in the same professional activities and allocate their time among academic pursuits in about the same ways. The mere performance of activities associated with teaching does not assure that the instruction is effective.

3. Eighty-five items are listed that characterize best teachers as perceived by students, and 54 items are listed that characterize best teachers as perceived by colleagues. All items statistically discriminate best from worst teachers with a high level of significance.

4. Analysis of the items characterizing best teachers as perceived by students produced five scales, or components of effective performance.

(Table 1). Our conceptual interpretations of the scales are indicated by the headings assigned:

1) Analytic/Synthetic Approach
2) Organization/Clarity
3) Instructor-Group Interaction
4) Instructor-Individual Student Interaction
5) Dynamism/Enthusiasm

5. Analysis of the items characterizing best teachers as perceived by colleagues produced five scales (Table 2). They are described as follows:

1) Research Activity and Recognition
2) Intellectual Breadth
3) Participation in the Academic Community
4) Relations with Students
5) Concern for Teaching

We believe that the colleague scales are somewhat less useful and valid than the student scales as a basis for evaluating teachers.

6. The student scales were derived from a 1967 survey. A single summary description was phrased to express the nature of the component of effective teaching identified by the items composing each scale. Respondents to the 1968 survey rated their teachers on each of the five summary descriptions and also on each of the items from which the scales had been derived. Correlations of mean scores on the summary descriptions with mean scores on the full lists of respective items were very high. Thus, the five summary descriptions provide the basis for a short evaluation form demonstrated to be broad and highly discriminating.

7. In general, student ratings of best teachers showed only negligible correlations with academic rank of instructor, class level, number of courses previously taken in the same department, class size, required versus optional course, course in major or not, sex of respondent, class level of respondent, grade-point average, and expected grade in course.

8. A disproportionate number of best teachers were teaching seminar rather than lecture courses, and a wide range of excellence was revealed in the teaching of different subject areas.

9. Seventeen items describing the college goals of students were sorted into three scales:

1) Upward Mobility/Security
2) Self-Knowledge/Humanism
3) Career/Subject Mastery

10. Thirteen items describing objectives of teaching as perceived by students were sorted into two scales:

1) Contribution to General Development
2) Transmission of Fundamentals

11. Students evaluated the positive contributions made to their lives by best teachers in six areas: knowledge imparted, counsel given, objectives clarified, values developed, incentive elicited, and skills developed. Correlation of mean scores for these areas with mean scores for the components of effective teaching and with overall ratings of effectiveness of teaching are high.

12. Nine types of effective teachers were identified by analyzing individual patterns of relatively high and low scores on the five components of effective teaching. Overall ratings of teachers having the various patterns correlate with certain course and student variables.

13. Teachers rated as excellent by some observers and as poor by others are less even in their performance of the five components of effective teaching than are best teachers.

OBJECTIVE

Our general objective is to enable the Academic Senate to provide guidance and incentive that can help to improve teaching.

One major aim is to identify and describe effective teaching so that instructors can be helped to improve, and so that the much neglected field of preparing graduate students for the teaching function of academic life may be benefited. Articles allegedly describing good teaching are numerous, and many are sound, but most either largely represent the subjective judgment of individuals and committees, or are based on studies using small samples in restricted circumstances. Faculty members feel that there are different ways of teaching well, and believe that they disagree somewhat in the assessment of teaching. Hence, many instructors remain skeptical of advice offered. Reliable characterization of effective teaching is needed.

The other major aim is to find more valid, reliable, and effective means of incorporating the evaluation of teaching into advancement procedures. We believe this to be the most important single requirement for the improvement of university teaching. The incentive thereby provided will encourage instructors to devote the study, time, and effort necessary to do their best, and the stature of teaching will increase.

Eighty-seven percent of 303 members of the Davis faculty (and 92% of similar samples at six diverse colleges and universities) stated in 1968 that effectiveness of teaching *should be* "Quite important," or "Very important" as a criterion for advancement. Only 24% of the Davis sample (and 38% of the total sample) stated that effectiveness of teaching actually *is* "Quite important," or "Very important." Seventy-one percent of the Davis faculty (and 72% of the total sample) stated that the campus should have a formal procedure for evaluating teaching.

The University of California's system of review for advancement is as good as we know, and since the revision, in Sept., 1969, of "Instructions to Appointment and Promotion Committees," greater efforts have been made to evaluate teaching. Nevertheless, procedures are largely unstandardized and hence do not adequately evaluate the teaching of the majority of the faculty. Consequently, here as elsewhere, "research and creative work" usually has outweighed quality of teaching as a criterion for advancement.

At Davis, as at most institutions, the dossier furnished by the department chairman to support promotion has been of the utmost importance, yet there are inherent weaknesses in a system that places great weight on evaluations of teaching as traditionally prepared by chairmen (or deans): A chairman may himself be doubtfully qualified as a judge of teaching. Opinions solicited from his staff may be biased, may not constitute an adequate sample, and often are in part secondhand. Most available measures of involvement in teaching (e.g., number of courses taught, enrollments, number of advisees) do not necessarily correlate with quality of instruction. Classroom visitations are resisted or resented by most teachers, and hence are seldom utilized although considered by many administrators to be the most important element in evaluation. If a department is large, the chairman cannot visit any class more than once or twice, which is enough to judge certain elements of effective teaching but insufficient to make a comprehensive judgment. Classroom instruction is only part of the teaching function.

We believe that promotion letters cannot be improved sufficiently to achieve our objective unless new procedures assure that they include more thorough, more objective, and more comparable evaluations of teaching than have been usual in the past.

DEVELOPMENT OF TEACHER-DESCRIPTION SCALES

Collection of Data

Three questionnaires were distributed on the Davis campus in May, 1967, and one in May, 1968. A random sample of all students was asked to complete the first questionnaire. The 278 undergraduate and 60 graduate students who responded constituted 4% of the student body and 38% of those approached. Bias may have been introduced by self-selection. However, the sample was evenly divided between the sexes, it did not differ significantly from the population in regard to distribution by class level or major area (except perhaps for moderate over-representation in the humanities), and the mean of the overall grade point averages of students in the sample was identical to the grade point average of the population for the spring, 1967 quarter. Respondents supplied biographical information and their academic backgrounds, and answered questions on their college goals, on objectives of teaching, and on the teaching of instructors identified by them as the best and worst they had had in the previous year. Assurance was given that the identity of teachers would be kept in strict confidence, even from Davis members of the research team.

The second questionnaire was returned by 119 of the faculty, which was 54% of the random sample approached and 21% of the resident teaching faculty. The respondent was asked to identify a best and a worst among his colleagues and to answer, for each, questions about teaching activities observed outside the classroom, about in-class behavior, and about the presentation of talks and seminars.

The third questionnaire was returned by 162 members of the faculty who were not asked to complete the previous questionnaire. This was 80% of the random sample approached and 29% of the resident teaching faculty. Questions related to distribution of time among various academic pursuits.

Lastly, as a follow-up and validation study, a fourth questionnaire was distributed in 1968 to all students in 51 classes. These were selected to include in about equal numbers, those of instructors identified in 1967 as best teachers by three or more students or colleagues, those of instructors identified as worst teachers, and classes of another group not previously identified as either best or worst and presumed to include teachers of intermediate effectiveness. The 1015 respondents again provided biographical data, and answered questions on their college goals, on various objectives of teaching, and on the teaching of the given instructor. Overall ratings of the teachers were also secured.

Identification of Effective Teachers

It is important to know whether the various segments of the academic community concur in their identifications of the most effective and ineffective teachers.

We identified the instructors receiving three or more nominations as best teachers, and those receiving three or more nominations as worst teachers by the student respondents to our 1967 survey. In an independent study at Davis by Regan and Yonge, 57 of the same teachers were named by students as being particularly excellent or poor. Appendix A shows the very high degree of agreement between the two surveys: the chi square value indicates a level of significance of $p < .005$ (that is, fewer than 5 chances in 10,000 that the observed result is fortuitous). This result indicates that the two groups of students probably used closely similar criteria of judgment. Since the Yonge-Regan study had a 90% return, we consider this to be indirect evidence that self-selection by our respondents did not introduce significant bias into the designation by them of their best and worst teachers.

Further, in our 1968 survey, all students of 15 instructors named in 1967 by three or more students as best teachers, all students of 18 instructors similarly named previously as worst teachers, and all students of 18 instructors not previously nominated as best or worst teachers, rated the excellence of their instructors. Ratings were along a seven-point continuum from "Among the very worst" to "Among the very best." Differences among the mean scores for the best, not nominated, and worst teachers of the previous year were all significant well below the .01 level.

Finally, each of 119 faculty respondents identified colleagues they considered outstanding and poor teachers. Of those named, 66 were common to the 1967 student sample. Appendix B shows the very high agreement between the two groups; again, $p < .0005$.

Components of Effective Teaching

Many researchers have sorted individual items describing aspects of effective teaching into related groups, thus identifying basic components, dimensions, or scales of such teaching. Teacher-rating forms developed by students commonly do the same. Scales have been variously determined by subjective examination of a list of items or by factor analysis, which establishes mathematically the tendency of responses to the various items to associate in clusters. The number of scales developed in reports we have seen ranges from 4 to 13. Nevertheless, 4 to 5 particular scales (i.e., knowledge, presentation, relation with students, enthusiasm) appear rather consistently, even though the terminology differs. Our scales are generally consistent with those of previous studies.

Scales for student characterization of effective teaching were established by factor analysis of 91 items describing the teaching of 338 best teachers as identified by respondents to our 1967 survey. (Our method was a principal-components analysis with a varimax rotation.) After several analyses, a five-factor solution was selected as giving the maximum number of distinct and interpretable components of effective teaching.

Table 1 presents the five scales and the included items. The factor coefficients from the 1968 survey are listed. Coefficients of .40 and higher are generally considered good; our cutoff value is .43. No item appears in more than one scale. Our conceptual interpretations of the scales are as follows:

Scale 1, *Analytic/Synthetic Approach,* is scholarship, with emphasis on breadth, analytic ability, and conceptual understanding.

Scale 2, *Organization/Clarity,* is skill at presentation, but as subject-related, not student-related, and is not merely rhetorical skill.

Scale 3, *Instructor-Group Interaction,* is rapport with the class as a whole, sensitivity to class response, and skill at securing active class participation.

Scale 4, *Instructor-Individual Student Interaction,* is mutual respect and rapport between the instructor and the individual student.

Scale 5, *Dynamism/Enthusiasm,* is the flare and infectious enthusiasm that comes with confidence, excitement for the subject, and pleasure in teaching.

Responses describing the performance of worst teachers were also subjected to factor analysis, but the results did not provide readily interpretable scales. The items showed less consistent relationships than they did for best teachers. Thus, ineffective teachers were characterized best by their lack of attributes associated with effective teaching, rather than by possession of attributes of poor teaching.

Scales for the characterization of effective teachers by colleagues were prepared by factor analysis of 67 items describing the behavior of 84 best teachers identified by 119 members of the faculty.

Table 2 presents the five scales, which were established by the same method of factor analysis as for the student data. The factor coefficients of the included items are listed, the cutoff value being .37. Alpha reliabilities range from .65 to .86. Intercorrelations among the scales are low or negligible. Our conceptual interpretations of the scales are indicated by the headings assigned:

Scale 1. Research Activity and Recognition
Scale 2. Intellectual Breadth
Scale 3. Participation in the Academic Community
Scale 4. Relations with Students
Scale 5. Concern for Teaching

Utility of the Scales

The scales derived from the characterization of effective teaching by students provide conceptual understanding of the components of such teaching. Having been developed from items to which most students of a large random sample could respond, the student scales are applicable to most kinds of university-level teaching. Attention to the scales helps to assure that the major components of effective performance are considered when teaching or evaluating teaching. Many of the rating forms in our files (assembled from various campuses) fail in this regard.

In order to learn if we could develop an effective short evaluation form, we phrased a summary description for each of the student scales derived from the 1967 survey. Each description was intended to express the component of effective teaching identified by the items composing that scale. The 1968 survey asked respondents to rate their teachers on each of the five summary descriptions. It also repeated the full set of original items from which the scales had been established. Correlations of mean scores on the summary descriptions with mean scores on the full list of respective items (N = 51) were very high (coefficients ranging from .88 to .96. Thus, a short-form rating instrument is established that is quickly answered, yet is objectively known to be broad, balanced, and highly discriminating between effective and ineffective teachers.

The five summary descriptions that we recommend follow. We suggest that if these scales are used for evaluations, the respondent be asked to use a seven-point continuum ranging from "Low score" to "High score."

1. Has command of the subject, presents material in an analytic way, contrasts various points of view, discusses current developments, and relates topics to other areas of knowledge.

2. Makes himself clear, states objectives, summarizes major points, presents material in an organized manner, and provides emphasis.

3. Is sensitive to the response of the class, encourages student participation, and welcomes questions and discussion.

4. Is available to and friendly toward students, is interested in students as individuals, is himself respected as a person, and is valued for advice not directly related to the course.

5. Enjoys teaching, is enthusiastic about his subject, makes the course exciting, and has self-confidence.

For reasons explained later in this report, we believe that colleague Scale 5, *Concern for Teaching,* is the most useful of the colleague scales. This is a summary description of that scale: *Expresses interest and concern in the quality of his teaching, discusses teaching in general with his colleagues, seeks advice regarding his teaching, and is sought by others for counsel on their teaching.*

Respondents to the 1968 student survey made a single overall rating of the effectiveness of their teachers on a continuum of 1 to 7. Scale 5, *Dynamism/Enthusiasm,* is the most highly related with the teachers named as best, and Scale 2, *Organization/Clarity,* is in second place. For all the correlations, $p < .001$.

[The above comprises about half of the full report. The sections on "Relation of Ratings of Teachers to the Course and the Student" and "Discussion" have not been included. The complete report is available from the Center for Research and Development in Higher Education, University of California at Berkeley.]

TABLE 1. COMPONENTS OF EFFECTIVE TEACHING AS PERCEIVED BY STUDENTS

SCALE 1. ANALYTIC/SYNTHETIC APPROACH

	Factor coefficient
1. Discusses points of view other than his own	.70
2. Contrasts implications of various theories	.66
3. Discusses recent developments in the field	.64
4. Presents origins of ideas and concepts	.60
5. Gives references for more interesting and involved points	.53
6. Presents facts and concepts from related fields	.53
7. Emphasizes conceptual understanding	.46

SCALE 2. ORGANIZATION/CLARITY

8. Explains clearly	.78
9. Is well prepared	.63
10. Gives lectures that are easy to outline	.62
11. Is careful and precise in answering questions	.61
12. Summarizes major points	.51
13. States objectives for each class session	.50
14. Identifies what he considers important	.47

SCALE 3. INSTRUCTOR-GROUP INTERACTION

15. Encourages class discussion	.70
16. Invites students to share their knowledge and experiences	.65
17. Clarifies thinking by identifying reasons for questions	.64
18. Invites criticism of his own ideas	.62
19. Knows if the class is understanding him or not	.58
20. Knows when students are bored or confused	.57
21. Has interest and concern in the quality of his teaching	.48
22. Has students apply concepts to demonstrate understanding	.43

SCALE 4. INSTRUCTOR-INDIVIDUAL STUDENT INTERACTION

23. Has a genuine interest in students	.74
24. Is friendly toward students	.71
25. Relates to students as individuals	.69
26. Recognizes and greets students out of class	.68
27. Is accessible to students out of class	.65
28. Is valued for advice not directly related to the course	.64
29. Respects students as persons	.60

SCALE 5. DYNAMISM/ENTHUSIASM

30. Is a dynamic and energetic person	.80
31. Has an interesting style of presentation	.76
32. Seems to enjoy teaching	.74
33. Is enthusiastic about his subject	.65
34. Seems to have self-confidence	.64
35. Varies the speed and tone of his voice	.63
36. Has a sense of humor	.53

Based on the 1968 survey. N = 1015

TABLE 2. COMPONENTS OF THE ACTIVITIES OF EFFECTIVE TEACHERS AS PERCEIVED BY COLLEAGUES

SCALE 1. RESEARCH ACTIVITY AND RECOGNITION

	Factor coefficient
1. Does work that receives serious attention from others	.69
2. Corresponds with others about his research	.69
3. Does original and creative work	.64
4. Expresses interest in the research of his colleagues	.55
5. Gives many papers at conferences	.55
6. Keeps current with developments in his field	.49
7. Has done work to which I refer in teaching	.48
8. Has talked with me about his research	.38

SCALE 2. INTELLECTUAL BREADTH

9. Seems well read beyond the subject he teaches	.66
10. Is sought by others for advice on research	.60
11. Can suggest reading in any area of his general field	.59
12. Knows about developments in fields other than his own	.51
13. Is sought by colleagues for advice on academic matters	.43

SCALE 3. PARTICIPATION IN THE ACADEMIC COMMUNITY

14. Encourages students to talk with him on matters of concern	.60
15. Is involved in campus activities that affect students	.58
16. Attends many lectures and other events on campus	.47
17. Has a congenial relationship with colleagues	.39

SCALE 4. RELATIONS WITH STUDENTS

18. Meets with students informally out of class	.58
19. Is conscientious about keeping appointments with students	.57
20. Meets with students out of regular office hours	.57
21. Encourages students to talk with him on matters of concern	.55
22. Recognizes and greets students out of class	.37

SCALE 5. CONCERN FOR TEACHING

23. Seeks advice from others about the courses he teaches	.70
24. Discusses teaching in general with colleagues	.60
25. Does not seek close friendships with colleagues (Negative)	—.47
26. Is someone with whom I have discussed my teaching	.45
27. Is interested in and informed about the work of colleagues	.44
28. Expresses interest and concern about the quality of his teaching	.40

N = 119

A SUGGESTED FORM FOR STUDENT EVALUATION OF TEACHING

STUDENT EVALUATION OF TEACHING — UNIVERSITY OF CALIFORNIA AT DAVIS

Instructor .... Department .... Course number ....
Year .... Quarter .... Teaching observed (give approximate number of each):
Lectures .... Labs ....
Conference-Discussion-Seminar .... Other (specify) ....

A. Each of these statements describes a basic component of teaching. Give the instructor an overall rating for each component, reserving the highest scores for unusually effective performance.

	Low Score						*High Score*
1. Has command of the subject, presents material in an analytic way, contrasts points of view, discusses current developments, and relates topics to other areas of knowledge.	1	2	3	4	5	6	7
2. Makes himself clear, states objectives, summarizes major points, presents material in an organized manner, and provides emphasis.	1	2	3	4	5	6	7
3. Is sensitive to the response of the class, encourages student participation, and welcomes questions and discussion.	1	2	3	4	5	6	7
4. Is available to and friendly toward students, is interested in students as individuals, is himself respected as a person and is valued for advice not directly related to the course.	1	2	3	4	5	6	7
Enjoys teaching, is enthusiastic about his subject, makes the course exciting, and has self-confidence.	1	2	3	4	5	6	7

B. These items are not covered in the statements above and thus extend the evaluation.

	Low Score				*High Score*	*Doesn't apply or don't know*
6. Has increased my appreciation for the subject.	1	2	3	4	5	☐
7. Keeps well informed about the progress of the class.	1	2	3	4	5	☐
8. Anticipates problems and makes difficult topics easy to understand.	1	2	3	4	5	☐
9. Is an excellent speaker.	1	2	3	4	5	☐
10. Quickly grasps what a student is asking or telling him.	1	2	3	4	5	☐
11. Presents the aesthetic and emotional values of the subject.	1	2	3	4	5	☐
12. Relates class topics to students' lives and experience.	1	2	3	4	5	☐
13. Gives interesting and stimulating assignments.	1	2	3	4	5	☐
14. Gives examinations that require creative, original thinking.	1	2	3	4	5	☐
15. Gives examinations that have instructional value.	1	2	3	4	5	☐

C. Additional items may be presented by the instructor and/or department.

D. You are invited to comment further on the course and/or effectiveness of the instruction.

Bibliography

The following bibliography has been divided into two sections. The first section emphasizes research studies dealing specifically with the evaluation of college teaching but includes important materials more broadly concerned with the recognition of effective teaching. The second section adds other research studies, including doctoral dissertations, and a range of material of a more general nature. In both lists, some materials chiefly concerned with the evaluation of elementary and secondary teachers have been included, but only those which are relevant to college and university teaching. Both lists are selective, particularly with regard to studies and articles prior to 1950. A number of individual institutional studies are included, but no attempt has been made to locate all such materials. Grateful acknowledgement is made to David S. Holmes, psychology department, University of Texas, for making available his extensive annotated bibliography on this subject.

I

Abramson, Paul. (ed.) "Sauce for the Goose is Sauce for the Gander," *College Management,* I:27-31, August, 1966.

Anderson, John E. "The Auto-Critique Method of Instructional Evaluation," *Proceedings of the Fourth Junior College Administrative Teams Institute,* Pensacola, Florida, July, 1964.

Anikeeff, A. M. "Factors Affecting Students' Evaluation of College Faculty Members," *Journal of Applied Psychology,* 37:458-60, 1953.

Balyeat, Ralph E. "Factors Related to the Promotion of University Personnel," (*Cooperative Research Project,* S-017), Athens:University of Georgia, 1964-5, (multilith).

Bannister, John, *et al.* "Evaluating College Teaching," *Curriculum Reporter Supplement No. 1,* San Jose State College, December, 1961.

Bendig, A. W. "An Inverted Factor Analysis Study of Student-Rated Introductory Psychology Instructors," *Journal of Experimental Education,* 21:333-36, 1953.

.... "Student Achievement in Introductory Psychology and Student Ratings of the Competence and Empathy of Their Instructors," *Journal of Psychology,* 36:427-33, 1953.

.... "The Relation of Level of Course Achievement to Student's Instructor and Course Ratings in Introductory Psychology," *Educational and Psychological Measurement,* 13:437-48, 1953.

Bernhard, H. A. "A Survey of Faculty-Evaluation Programs Administered by Chapters of Tau Beta Pi," *School and Society,* 72:102-05, 1950.

Blackman, A. F., *et. al.* "Students Rate Their Profs and Courses," *Phi Delta Kappan,* 48:266-69, February, 1967.

Bousfield, W. A. "Students' Ratings of Qualities Considered Desirable in College Professors," *School and Society,* 51:253-56, 1940.

Boyer, Marcia. "Teacher Evaluation: Toward Improving Instruction," *ERIC Junior College Research Review,* 4:1-4, January, 1970.

Brickman, W. W. "Student Evaluation of Professors," (editorial), *School and Society,* 94:143-44, 1966.

Brown, D. R., *et al. Report of the Committee on Student Evaluation of Courses,* Ann Arbor:University of Michigan, November, 1968.

Bryan, R. C. "Teacher's Image Is Stubbornly Stable," *Clearing House,* 40:459-61, 1966.

Clark, K. W., and Keller, R. J. "Student Ratings of College Teaching," *in* Ruth Eckert and R. J. Keller (eds.) *A University Looks at its Program,* Minneapolis: University of Minnesota, 197-212, 1954.

Coffman, Wm. E. "Determining Students' Concepts of Effective Teaching from Their Ratings of Instructors," *Journal of Educational Psychology,* 45:227-86. 1954.

Cohen, Arthur M., and Brawer, Florence B. *Measuring Faculty Performance,* Monograph No. 4., Washington, D.C.:American Association of Junior Colleges, 1969.

Cole, Luella. *The Background for College Teaching,* New York:Farrar and Rinehart, 572-89, 1940.

Cosgrove, D. J. "Diagnostic Rating of Teacher Performance," *Journal of Educational Psychology,* 50:200-04, 1959.

Crannell, C. W. "A Preliminary Attempt to Identify the Factors in Student-Instrutor Evaluation, *Journal of Psychology,* 36:417-22, 1953.

Cronkhite, Bernice Brown. *A Handbook for College Teachers,* Cambridge, Massachusetts:Harvard, 1950.

Crawford, P. L., and Bradshaw, H. L. "Perception of Characteristics of Effective University Teachers: A Scaling Analysis," *Educational and Psychological Measurement,* 28:1079-85, 1968.

Deshpande, A. S., and Webb, S.C. "Student Perceptions of Instructor Teaching Goals. III. Internal Structure of Ratings," *(Research Memorandum,* 68-5), Office of Evaluation Studies, Georgia Institute of Technology, 1968.

Detchen, L. "Shall the Student Rate the Professor?" *Journal of Higher Education,* 11:146-54, 1940.

Dick, W. *Course Attitude Questionnaire: Its Development, Uses and Research Results,* University Park:Division of Instructional Services, Pennsylvania State University, 1967.

Downie, N. M. "Student Evaluation of Faculty," *Journal of Higher Education,* 23: 495-96, 503, December, 1952.

Dressel, P. L. "The Current Status of Research on College and University Teaching," in *The Appraisal of Teaching in Large Universities,"* Ann Arbor:University of Michigan, 7-16, 1959.

Drucker, A. J., and Remmers, H. H. "Do Alumni and Students Differ in Their Attitudes Toward Instructors?" *Journal of Educational Psychology,* 42:129-43, March, 1951.

Dwyer, F. M. *A Review of Characteristics and Relationships of Selected Criteria for Evaluating Teacher Effectiveness,* University Park:Division of Instructional Services, Pennsylvania State University, 1968.

Eckert, R. E. "Ways of Evaluating College Teaching," *School and Society,* 71:65-9, 1950.

...., and Neale, D.C. "Teachers and Teaching," *Review of Educational Research,* 1965, 35:307-08, October, 1965.

Elliott, D. N. "Characteristics and Relationships of Various Criteria of College and University Teaching," *Purdue University Studies in Higher Education,* 70:5-61, 1950.

Ericksen, Stanford C. "Evaluation of Instruction by Students," *Memo to the Faculty,* Ann Arbor:University of Michigan, 10:1-4, February, 1965.

Estrin, H. A., and Godwin, L. R. "Evaluating English Teaching," *Improving College and University Teaching,* 10:193-96, Autumn, 1962.

...., and Goode, D. M. (eds.) *College and University Teaching,* Dubuque, Iowa: Wm. C. Brown Company, 1964.

Fitch, Naomi. *Evaluation of Instructors in California Junior Colleges,* Berkeley:University of California, 1965.

French, G. M. *College Students' Concept of Effective Teaching Determined by an Analysis of Teacher Ratings,* Seattle:University of Washington, 1957.

Gadzella, B. M. "College Students' Views and Ratings of an Ideal Professor," *College and University,* 44:89-96, Fall, 1968.

Gage, N. L. "The Appraisal of College Teaching: An Analysis of Ends and Means," *Journal of Higher Education,* 32:17-22, January, 1961.

.... "Ends and Means in Appraising College Teaching," in *The Appraisal of Teaching in Large Universities,* Ann Arbor:University of Michigan, 58-64, 1959.

Garverick, C. M., and Carter, H. D. "Instructor Ratings and Expected Grades," *California Journal of Educational Research,* 13:218-21, 1962.

Gates, M. and Collins, W. B. "Students' Perceptions of Professors," *Improving College and University Teaching,* 17:234-36, Autumn, 1969.

Gibb, C. A. "Classroom Behavior of the College Teacher," *Educational and Psychological Measurement,* 15:254-63, 1955.

Goode, Delmer M. (ed) "Encouraging and Appraising Teaching Quality," contains articles on evaluation by Brogan, Falk, Whitlock and Montgomery, Bryan, Daugherty, and Neeley, *Improving College and University Teaching,* 16:191-209, Summer, 1968.

Goodhartz, Abraham S. "Student Attitudes and Opinions Relating to Teaching at Brooklyn College," *School and Society,* 68:345-49, 1948.

Gustad, John W. "Policies and Practices in Faculty Evaluation," *Educational Record,* 42:194-211, 1961.

Guthrie, Edwin R. "Evaluation of Faculty Service," *AAUP Bulletin,* 31:255-62, Summer, 1945.

.... "The Evaluation of Teaching," *Educational Record,* 30:109-15, 1949.

.... "The Evaluation of Teaching: A Progress Report," Seattle:University of Washington, 1954 (lithoprint).

Hayes, R. B. "A Way to Measure Classroom Teaching Effectiveness," *Journal of Teacher Education,* 14:168-76, 1963.

Harris, T. L., Kiefert, J. J., and Darby, M. D. "Attitudes Expressed by Students Toward a Beginning Course in Educational Psychology," *Journal of Educational Research,* 62:344-50, 1969.

Heilman, J. D. and Armentrout, W. D. "The Ratings of College Teachers on Ten Traits by Their Students," *Journal of Educational Psychology,* 27:197-216, 1936.

Hildebrand, M., and Wilson, R. C. *Report to the Faculty on Effective University Teaching and Its Evaluation,* University of California at Davis (multilith), 1970.

Hodgson, T. F. *The General and Primary Factors in Student Evaluation of Teaching Ability,* Seattle:University of Washington (mimeo), 1958.

Hoyt, Donald P. "Identifying Effective Classroom Procedures," *Research Report* No. 7, Manhattan:Kansas State University, 1969.

.... *Improving Instruction Through Student Feedback,* Office of Educational Research, Manhattan:Kansas State University, 1969.

.... "Measurement of Effectiveness, Instructional Effectiveness" *Research Report* No. 6, Manhattan:Kansas State University, 1969.

Hudelson, Earl, "The Validity of Student Rating of Instructors," *School and Society,* 73:265-66, April 28, 1951.

Hussain, K. M., and Leestamper, R. "Survey on Criteria of Teaching Effectiveness at New Mexico State University," Las Cruces:New Mexico State University, 1968.

Isaacson, Robert L., McKeachie, W. J., and Milholland, J. E. "Correlation of Teacher Personality Variables and Student Ratings," *Journal of Educational Psychology,* 54:110-17, 1963.

............, *et al.* "Dimensions of Student Evaluations of Teaching," *Journal of Educational Psychology,* 55:344-51, December, 1964.

Jensen, Paul H. *A Study of Self-Evaluation Applied to In-Service Education,* including annotated bibliography, Monmouth:Teaching Research Division, Oregon State System of Higher Education, 1968.

Karman, Thomas A. "Faculty Evaluation," *Liberal Education,* 55:539-44, 1969.

Kilpatrick, Gordon. *Another Look at Teacher Evaluation in American Junior Colleges,* El Camino College, California, 1967.

King, James H., and Ellis, Joseph R. *A National Survey of Factors Judged to be Associated with Ineffectiveness of College Educators,* DeKalb:Northern IllinoisUniversity, (mimeo), 1969.

Langen, T. D. "Student Assessment of Teaching Effectiveness," *Improving College and University Teaching,* 14:22-25, Winter, 1966.

Lee, Calvin B. (ed.) *Improving College Teaching,* Part 5, "The Evaluation of Teaching Performance," Washington, D.C.:The American Council on Education, 259-343, 1967.

Leftwich, Wm. H., and Remmers, Herman H. "A Comparison of Graphic and Forced-Choice Ratings of Teaching Performance at the College and University Level," *Purdue University Studies in Higher Education,* XCII, 1-35, 1962.

Lovell, G. D., and Haner, C. F. "Forced-Choice Applied to College Faculty Rating," *Educational and Psychological Measurement,* 15:291-304, 1955.

Luthans, Fred. *The Faculty Promotion Process: An Empirical Analysis of the Administration of Large State Universities,* Bureau of Business and Economics Research; Iowa City:University of Iowa, 1967.

McDonagh, Edward. "Let's Grade the Professors," *AAUP Bulletin,* 30:83-6, February, 1944.

McGrath, Earl J. "Characteristics of Outstanding College Teachers," *Journal of Higher Education,* 148-52, March, 1962.

McKeachie, W. J. (chairman) "Conference on Appraisal of Teaching in Large Universities," *The Appraisal of Teaching in Large Universities,* Ann Arbor:University of Michigan, 1959.

.... "The Improvement of Instruction," *Review of Educational Research,* 30: 351-61, October, 1960.

.... "Procedures and Techniques of Teaching: A Survey of Experimental Studies," *in* (N. Sanford, ed.) *The American College,* New York:John Wiley, 1962, basis for "Current Research on Teaching Effectiveness," *Improving College and University Teaching,* 10:15-19, Winter, 1962.

.... "Student Ratings of Faculty," *AAUP Bulletin, 55:439-44,* December, 1969.

............, *et al. Research on the Characteristics of Effective College Teaching,* U. S. Department of Health, Education, and Welfare, Office of Education Project No. 850, Ann Arbor:University of Michigan, 1964.

............, and Solomon, D. "Student Ratings of Instructors: A Validity Study," *Journal of Educational Research,* 51:379-82, 1958.

.... "Teaching Tips," Ann Arbor: George Wahr Publishing Company, 5th Edition, 1965, D. C. Heath, 1969.

McNeil, J. D. "Concomitants of Using Behavioral Objectives in the Assessment of Teaching Effectiveness," *Journal of Experimental Education,* 36:69-74, 1967.

Maslow, A. H., and Zimmerman, W. "College Teaching Ability, Scholarly Activity, and Personality," *Journal of Educational Psychology,* 47:185-89, 1956.

Mayhew, L. B., and Samalonis, B., Renner, R. R., Spaights, E., and Freehill, M. F. "Appraisal: The Pointing Finger," (five articles on evaluation), *Improving College and University Teaching,* 15:5-19, Winter, 1967.

Meredith, Gerald M. "Dimensions of Faculty-Course Evaluation," *The Journal of Psychology,* 73:27-32, 1969.

Moore, J. E. "A Further Study of the Annoying Habits of College Professors," *Journal of Abnormal Social Psychology,* 32:368-75, 1937.

Morsh, J. E., Burgess, G. G., and Smith, P. N. "Student Achievement as a Measure of Instructor Effectiveness," *USAF Personnel and Training Research Center Development Report,* 1955.

Morsh, J. E., and Wilder, Eleanor W. "Identifying the Effective Instructor: A Review of the Quantitative Studies 1900-1952," *Research Bulletin* AFPTC-TR 54-44, Air Force Personnel and Training Research Center, 1954.

Mueller, Francis J. "Trends in Student Ratings of Faculty," *AAUP Bulletin,* 37:319-24, Summer, 1951.

Musella, D., and Rusch, R. "Student Opinion on College Teaching," *Improving College and University Teaching,* 16:137, Spring, 1968.

Overturf, C. L., and Price, E. C. *Student Ratings of Faculty at St. Johns River Junior College,* Palatka, Florida: St. Johns River Junior College, 1966.

Perry, R. A. "Evaluation of Teaching Behavior Seeks to Measure Effectiveness," *College and University Business,* 47:18, October, 1969.

Rayder, N. F. "College Student Ratings of Instructors," *Journal of Experimental Education,* 37:76-81, 1968.

Reed, Anna, Maguire, Frank, Thomson, Wm., and Van der Vate, Harriett. *The Effective and the Ineffective College Teacher,* New York:American Book Company, 1935.

Remmers, H. H. "Appraisal of College Teaching Through Ratings and Student Opinion," *27th Year Book American Association of Colleges of Teacher Education,* Chicago:University of Chicago Press, 227-40, 1939.

............, and Baker, P. C. *Attitude Toward Student Ratings of Instruction,* Lafayette, Indiana:Division of Educational Reference, Purdue University, 1950.

............, Davenport, K. S., and Potter, A. A. "The Best and the Worst Teachers of Engineering," *Studies in Higher Education LVII, Bulletin of Purdue University,* 1-20, October, 1946.

.... "The College Professor as the Student Sees Him," *Purdue University Studies in Higher Education XI,* 1-63, 1929.

............,Elliott, D. N., and Hobson, R. L. "Norms and Illustrative Profiles of Teachers and Administrators," *Report Number 1, Indiana College and University Staff Evaluation,* Lafayette, Indiana: Purdue Research Foundation, 1948.

............, Martin, F. D., and Elliott, D. N. "Are Student Ratings of Their Instructors Related to Their Grades?" *Purdue University Studies in Higher Education LXVI,* 17-26, 1949.

.... "On Students' Perceptions of Teachers' Effectiveness," in *The Appraisal of Teaching in Large Universities,* Ann Arbor: University of Michigan, 1959.

.... "Rating Methods in Research on Teaching," *Handbook of Research on Teaching* (N. L. Gage, ed.), Chicago: Rand McNally and Company, 329-78, 1963.

.... "Reliability and Halo Effect of High School and College Students' Judgments of Their Teachers," *Journal of Applied Psychology,* 18:619-30, 1934.

.... "To What Extent do Grades Influence Student Ratings of Instructors?" *Journal of Educational Research,* 21:314-16, 1930.

.... (ed.) *Studies in College and University Staff Evaluation No. XV, Studies in Higher Education LXX, Bulletin of Purdue University,* 1-92, March, 1950.

Report of the Committee on College and University Teaching, American Association of University Professors, 1933.

Report to the University, "Teaching, Research and the Faculty," *The Study of Education at Stanford,* 8:72-144, 1969.

Riley, J. W., Jr., Ryan, B. F., and Lifshitz, M. *The Student Looks at His Teacher,* New Brunswick, New Jersey: Rutgers University Press, 1950.

Rohrer, Robert, and Wertheim, Richard. "In Spite of Brickbats and Bugs," *Emory College Today,* 42:20-22, Summer, 1967.

Rothwell, C. Easton, *et al. The Importance of Teaching: A Memorandum to the New College Teacher,* New Haven, Connecticut: The Hazen Foundation, 1968.

Rowland, R. "Can Teaching Be Measured Effectively?" *Improving College and University Teaching,* 18:153-57, Summer, 1970.

Russell, H. E., and Bendig, A. W. "An Investigation of the Relation of Student Ratings of Psychology Instructors to Their Course Achievement When Academic Aptitude is Controlled," *Educational and Psychological Measurement,* 13-626-35, 1953.

Ryans, David G. *Characteristics of Teachers,* Washington, D.C.: American Council on Education, 1960.

.... "Research on Teacher Behavior in the Context of the Teacher Characteristics Study," *Contemporary Research on Teacher Effectiveness* (B. J. Biddle and W. J. Ellena, eds.), New York: Holt, Rinehart and Winston, 67-101, 1964.

Schilling, Bernard N. "De Rerum Natura: Supplement to Report of the Committee . . ." *AAUP Bulletin,* 42:445-63, Autumn, 1956.

Schneider, Franz. *More Than an Academic Question,* Berkeley, California: Pestalozzi Press, 1945.

.... *Students Examine Their Professors,* Berkeley, California: Pestalozzi Press, 1939.

Shao, Otis H. "Faculty Evaluation Plan in a Liberal Arts College," *Liberal Education,* 54:413-17, October, 1968.

Simpson, Ray H., and Brown, E. S. "College Learning and Teaching," *University of Illinois Bulletin,* Urbana, Illinois, 49:75, June, 1952.

Slobin, D. Y., and Nichols, D. G. "Student Rating of Teaching," *Improving College and University Teaching,* 17:244-47, Autumn, 1969.

Smith, G. Kerry (ed.) *Current Issues in Higher Education,* Washington, D.C., Association for Higher Education, 1966: John Centra, "Evaluating Teaching," Ralph Balyeat, "Institutional Practice Concerning Faculty Status and Reward," Lloyd S. Woodburne "Guidelines for Student Ratings of Faculty;" 1967: Louis Long, "Why Evaluate?" Philip Werdell, "Student Participation in Effective Program of Faculty Evaluation."

Smith, Page. "Teaching, Research, and Publications as They Affect Academic Performance and Promotion," *Journal of Higher Education,* 32:199-205, 1961.

Solomon, D., Rosenberg L., and Bezdek, W. E. "Teacher Behavior and Student Learning," *Journal of Educational Psychology,* 55:23-30, 1964.

Solomon, D. "Teacher Behavior Dimensions, Course Characteristics, and Student Evaluations of Teachers," *American Educational Research Journal,* 3:35-47, 1966.

Starrack, J. A. "Student Rating of Instruction," *Journal of Higher Education,* 5:88-90, February, 1934.

Stecklein, John E. "Colleges and Universities: Appraisal of College Teachers," *Encyclopedia of Educational Research,* New York: Macmillan Company, 268-88, 1960.

Stewart, C. T., and Malpass, L. F. "Estimates of Achievement and Ratings of Instructors," *Journal of Educational Research,* 59:347-50, 1966.

Stuit, Dewey, and Ebel, Robert. "Faculty Evaluation: Instructor Rating at a Large State University," *Special Reports,* No. 3785, University of Iowa, Iowa City.

Taft, Marvin I. "A Systematic Method for Evaluating Teachers," *Journal of Engineering Education,* 59:852-57, March, 1969.

Taylor, W. S., Hankins, F. H., and Lazerowitz, A. A. "Students' Ratings of Instruction; A Study Made at Smith College," *Journal of Higher Education,* 18-202-6, 1947.

Tyler, Ralph C. "The Evaluation of Teaching," in *The Two Ends of the Log,* (Russell Cooper, ed.) Minneapolis: University of Minnesota, 1958.

.... "The Evaluation of Teaching," in *Preparing College Teachers* (Albright, A. D., and Barrows, J. E., eds.), Lexington: University of Kentucky, 1959.

Voeks, V. W., and French, G. M. "Are Student Ratings of Teachers Affected by Grades?" *Journal of Higher Education,* 31:330-34, June, 1960.

Waschler, N. F. "Instructional and Course Rating by Students," (Summary of Report by Committee to Study Instructional Rating, Wharton School of Finance and Commerce, University of Pennsylvania), *Educational Record, 32:89-95,* 1951.

Weaver, C. H. "Instructor Rating by College Students," *Journal of Educational Psychology,* 51:22-25, 1960.

Weber, C. A. "Some Characteristics of College Teachers," *Journal of Educational Research,* 46:685-92, 1953.

Werdell, Philip. *Course and Teacher Evaluation,* 2115 S Street, N.W., Washington, D.C.: U. S. National Student Association, January, 1966, Second Printing, August, 1967.

Yamamoto, K., and Dizney, H. F. "Eight Professors — A Study on College Students' Preferences Among Their Teachers," *Journal of Educational Psychology,* 57:146-50, 1966.

II

Allen, Dwight W., and Gross, R. E. "Microteaching," *NEA Journal,* 55:25-6, December, 1965.

Amidon, E., and Simon, A. "Teacher-Pupil Interaction," *Review of Educational Research,* 35:130-39, April, 1965.

Anderson, C. C., and Hunka, S. M. "Teacher Evaluation: Some Problems and a Proposal," *Harvard Educational Review,* 33 (1):74-95, 1963.

Anderson, Carl Leonard. "The Student Looks at His Learning," *Improving College and University Teaching,* 2:65-6, November, 1954.

Anderson, Duane. *Report of Committee on Evaluation of Instruction,* American Association of Junior Colleges, March, 1970.

Arden, E. "Faculty as Teachers," *Educational Forum,* 32:447, May, 1968.

Arrowsmith, Wm. "The Future of Teaching," in *Improving College Teaching,* (C. B. Lee, ed.), Washington, D.C.: American Council on Education, 57-71, 1967.

Astin, A. W. "Classroom Environments in Different Fields of Study," *Journal of Educational Psychology,* 56:275-82, October, 1965.

Bard, Harry. "Teaching at the Junior College Level: Some Guideposts for the Improvement of Instruction," *Junior College Journal,* 437-40, April, 1962.

Barnard, Ellsworth. "More Thoughts on Grading the Professors," *AAUP Bulletin,* 31:263-74, Summer, 1945.

Barr, A. S. "An Expert Answers Questions Concerning Teacher Evaluation," *Illinois Education,* 49:9, May, 1961.

.... "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations," *Journal of Experimental Education,* 30:5-156, 1961.

Bayley, D. H. "Making College Teaching a Profession," *Improving College and University Teaching,* 15:115-19, Spring, 1969.

Boeck, C. H. "Dinosaurs and Dodo Birds in Higher Education," *Science Education,* 52:23-28, 1968.

Borenstein, W. "Reptiles and Reactions: Student Appraisal of Professors," *Educational Record,* 49:166, Spring, 1968.

Brandis, Royall. "The Rehabilitation of University Undergraduate Teaching," *Educational Record,* 45:56-63, Winter, 1964.

Brick, Michael, and McGrath, Earl J. *Innovation in Liberal Arts Colleges,* Institute of Higher Education, Teachers College Press, Columbia University, 1969.

Brock, E. H. "Rating Teaching Personnel for Advancement Purposes," *Journal of Engineering Education,* 44:593-99, June, 1954.

Broudy, Harry S. "Can We Define Good Teaching?" *Teachers College Record,* 70:583-92, April, 1969.

Carpenter, F., Van Egmond, E., and Jochem, J. "Student Preference of Instructor as a Function of Subject Matter," *Science Education,* 49:235-38, 1965.

Cobb, E. B. "Construction of a Forced-Choice University Instructor Rating Scale," (University of Tennessee), *Dissertation Abstracts,* 2202, 1956.

Cogan, M. L. "Research on the Behavior of Teachers: A New Phase," *Journal* of *Teacher Education,* 14:238-44, 1963.

Combs, Arthur W., and Mitzel, Harold E. "Can We Measure Good Teaching Objectively?" *NEA Journal, 53: 34-6,* January, 1964.

Cook, W. W., Leeds, C. H., and Collis, R. *Minnesota Teacher Attitude Inventory,* New York: Psychological Corporation, 1951.

Cortez: J. D. "The Design and Validation of an Evaluation Instrument to Appraise the Teaching Effectiveness of a College Instructor," (University of Denver), *Dissertation Abstracts,* 2435-A, 1967.

deBruin, H. C. "Quality Instruction," *Improving College and University Teaching,* 15:214-15, Autumn, 1967.

Dillman, T. "When College Students Grade the Faculty," *Today's Education,* 59:62+, February, 1970.

Domas, Simeon J., and Tiedeman, David V. "Teacher Competence: An Annotated Bibliography," *Journal of Experimental Education,* 19:101-218, 1950.

Dubin, R., and Taveggia, T. C. *The Teaching-Learning Paradox,* Eugene: University of Oregon, 1968.

Eells, W. C. *College Teachers and College Teaching,* an annotated bibliography, 1957, Atlanta, Georgia: Southern Regional Education Board; First Supplement, 1959; Second Supplement, 1962; Third Supplement by M. L. Litton and W. H. Stickler, 1967.

Flanagan, J. C. "The Critical Incident Technique," *Psychological Bulletin,* 51:327-58, July, 1954.

Flanders, N. A. *Teacher Influence, People, Attitudes and Achievement: Studies in Interaction Analysis,* Office of Education Cooperative Research Project 397, Minneapolis: University of Minnesota; 1960.

French, Sidney J., and Cooper, Russell M. *Pilot Project for Improving College Teaching — The Florida College Teaching Project,* Tampa: University of Southern Florida, Report #CRP-Z-001, June, 1967.

Frumkin, R. M. and Howell, W. H. "Effective College Teaching: A Preliminary Causal Analysis," *Journal of Educational Research,* 47:683-93, 1954.

Furniss, W. Todd. "Department and Faculty Profiles: An Aid to Judgment," *Liberal Education,* 49:354-65, October, 1963.

Garber, H. "Certain Factors Underlying the Relationship Between Course Grades and Student Judgments of College Teachers," (University of Connecticut), *Dissertation Abstracts,* 26 (11):6512, 1966.

Garrison, Lloyd L. "A Faculty Looks at Itself," *Improving College and University Teaching,* 2:63-4, November, 1954.

Getzels, J. W., and Jackson, P. W. "The Teacher's Personality and Characteristics," *Handbook of Research on Teaching,* (N. L. Gage, ed.), Chicago: Rand McNally and Company, Chapter 11, 1963.

Good, Carter V. *Teaching in College and University: A Survey of the Problems and Literature in Higher Education,* Baltimore: Warwick and York, 1929.

Goode, Delmer M. *These Books Were Stimulating: Seventy Books on College and University Teaching,* Corvallis: Oregon State College, 1955.

Gowin, D. B., and Payne, D. E. "Evaluating Instruction: Cross Perceptions of College Students and Teachers," *School Review,* 70:207-19, 1962.

Gray, Charles E. "The Teaching Model and Evaluation of Teaching Performance," *The Journal of Higher Education,* 40:636-42, 1969.

Gustad, John W. "On Improving College Teaching," *NEA Journal,* 53:37-8, March, 1964.

Hechinger, Fred M. "Now Grades for Teacher," *The New York Times,* Sunday, Section E, p. 9, October 24, 1965.

Heckman, Dale M., and Martin, Warren B. "Student-Faculty Interaction," *Inventory of Current Research on Higher Education, 1968,* New York: McGraw-Hill, 58-61, 1969.

Heiss, Ann M. "Berkeley Doctoral Students Appraise Their Academic Programs," *Educational Record,* 48:30-44, 1967.

Highet, Gilbert. *The Art of Teaching,* New York: Knopf, 1950.

Hildebrand, Joel H. "On Grading Without Judgment," *AAUP Bulletin,* 31:638-42, 1945.

Holland, J. B. "The Image of the Instructor as it is Related to Class Size," *Journal of Experimental Education,* 23:171-77, 1954.

Howsam, Robert B. *Who's A Good Teacher? Problems and Progress in Teacher Evaluation,* Burlingame: Joint Committee on Personnel Procedures of the California School Board Association and the California Teachers Association, 1960.

Jarrett, J. L. "The Self-Evaluating College Teacher," *Today's Education,* 58:40-1, January, 1969.

Jason, Hilliard. *Teaching and Research,* Commission on Undergraduate Education in the Biological Sciences, Publication No. 23, May, 1969.

.... "A Study of Medical Teaching Practices," *Journal of Medical Education,* 1258-84, December, 1962.

Jersild, A. T. "Characteristics of Teachers Who are 'Liked Best' and 'Disliked Most'," *Journal of Experimental Education,* 7:139-51, 1940.

Johnson, G. B. "An Evaluation Instrument for Analysis of Teacher Effectiveness," *Journal of Experimental Education,* 23:331-34, 1955.

Kahn, A. E. *Report of the Faculty Committee on the Quality of Undergraduate Instruction,* Unpublished Report, Ithaca, New York: Cornell University, 1965.

Kelly, F. J. *Improving College Instruction,* Washington, D.C.: American Council on Education, 1951.

Klapper, P. "The Professional Preparation of the College Teacher," *Journal of General Education,* 3:228-44, April, 1949.

Lacognata, A. A. "University Extension Faculty and Student Role Expectations: An Empirical Analysis," *Journal of Experimental Education,* 33:107-20, 1964.

Layton, David. *University Teaching in Transition,* Edinburgh: Oliver and Boyd, 1968.

Leese, Joseph. *Three Reports of the Project on the Improvement of College Teaching,* School of Education, SUNY at Albany, 1966.

Lewis, Edwin C. "An Investigation of Student-Teacher Interaction as a Determiner of Effective Teaching," *Journal of Educational Research,* 57:360-63, 1964.

Lewis, L. S. "Students' Images of Professors," *Educational Forum,* 32:185-90, January, 1968.

Long, H. B. "Teacher Behavior Most Disliked by Students," (Stanford) Ph.D. Dissertation, 1949.

Lynes, R. "The Teacher is an Iceberg," *Junior College Journal,* 36:8-12, April, 1966.

McGrath, E. "Superior Teaching in the Colleges of the Sixties," *Improving College and University Teaching,* 8:76-83, Summer, 1960.

McKeachie, W. J. "Psychology at Age 75: The Psychology Teacher Comes Into His Own," *American Psychologist,* 23:551-57, August, 1968.

.... "Research on Teaching at College and University Level," *Handbook of Research on Teaching,* (N. L. Gage, ed.), Chicago: Rand McNally and Company, 1118-72, 1963.

Mann, W. R. "Changes in the Level of Attitude Sophistication of College Students as a Means of Teacher Effectiveness," (University of Michigan), Dissertation Abstracts, 29: 2443-A, 1968.

Mathews, Kenneth P., and Milholland, John E. "Experience with a Post Hoc Procedure for Evaluating Methods of Teaching a Medical Subject," *Journal of Medical Education*, 229-39, March, 1960.

Medley, Donald M., and Mitzel, Harold E. "Measuring Classroom Behavior by Systematic Observation," *Handbook of Research on Teaching*, (N. L. Gage, ed.), Chicago: Rand McNally, 247-328, 1963.

Mitzel, Harold E. "Teacher Effectiveness," *Encyclopedia of Educational Research*, 3rd edition, Macmillan and Company, 1481-86, 1960.

O'Brien, J. C. "The Professor Stands Trial," *Educational Forum*, 34:39-45, November, 1969.

Payne, Fernandus, and Spieth, Evelyn Wilkinson. *An Open Letter to College Teachers*, Bloomington, Indiana: The Principia Press, Inc., 1935.

Perkins, D. "We Shall Gladly Teach," *The American Historical Review*, 62:291-309, January, 1957.

Peterson, Houston (ed.) *Great Teachers Portrayed by Those Who Studied Under Them*, New Brunswick, New Jersey: Rutgers University Press, 1946.

Pullias, Earl V., and Lockhart, Aileene, et al. *Toward Excellence in College Teaching*, Dubuque, Iowa: W. C. Brown, 1963.

Remmers, H. H. *Manual of Instructions for the Purdue Rating Scale for Instruction*. (Rev. ed.) Lafayette, Indiana: Purdue University Book Store, 1960.

Rogers, Carl. "The Facilitation of Significant Learning," *Instruction: Some Contemporary Viewpoints*, (L. Siegel, ed.), San Francisco: Chandler Publishing Company, 1967.

Ronan, W. W., and Prien, E. P. *Towards a Criterion Theory: A Review and Analysis of Research and Opinion*, Greensboro, North Carolina: Creativity Research Institute of the Richardson Foundation, Inc., 1966.

Resler, A. G. "The Influence of Needs Upon the Student's Perception of His Instructor," *Journal of Educational Research*, 58-282-86, 1965.

Ryans, David G. "Notes on the Rating of Teacher Performance," *Journal of Educational Research*, 47:695-705, 1954.

———, *et al.* "Report of the Committee on the Criteria of Teacher Effectiveness," *Review of Educational Research*, 22:238-63, 1952.

Sawyer, Broadus E. "Teacher Evaluation in the Liberal Arts College," *Journal of Teacher Education* 13:268-72, September, 1962.

Schmidt, Rita (ed.) *Insight: A View of the Faculty Through the Eyes of Their Students*, San Marcos, California: Palomar College, 1968.

Shane, H. G. "How Do They Rate You, Professor?" *NEA Journal*, 54:18-20, November, 1965.

Simpson, Ray H. "Self-Evaluation by Physical Science Instructors," *Science Education*, 50:58-64, 1966.

——— *Teacher Self-Evaluation*, New York: Macmillan, 1966.

Smalzried, N. T., and Remmers, H. H. "A Factor Analysis of the Purdue Rating Scale for Instructors," *Journal of Educational Psychology*, 34:363-67, September, 1943.

Smith, A. "College Teachers Evaluated by Students," *Sociology and Social Research*, 28:471-78, 1944.

Snedeker, J. H. "The Construction of a Forced-Choice Rating Scale for College Instructors, (University of Indiana), Dissertation Abstracts 1275, 1959.

Spencer, R. E. *The Illinois Course Evaluation Questionnaire: Manual of Interpretation*. Measurement and Research Division, Champaign: University of Illinois, 1968.

Trotter, Marilyn B. *Faculty-Student Relations: What Some Students Say About Faculty*, Student Counseling Center, Manhattan: Kansas State University, 1967.

Tuckman, B. W., and Oliver, W. F. "Effectiveness of Feedback to Teachers as a Function of Source," *Journal of Educational Psychology*, 59:297-301, 1968.

Vanderpool, Jeanette A. "Student Opinion — Sacred Cow or Booby Trap?" *Journal of Teacher Education*, 10:401-12, December, 1959.

Walsh, W. B. "Validity of Self-Report," *Journal of Counseling Psychology*, 14:18-23, 1967.

Watters, Wm. A. "Annotated Bibliography of Publications Related to Teacher Evaluation," *Journal of Experimental Education*, 22:351-67, 1954.

Webb, Clark, and Baird, J. Hugh. "Learning Differences Resulting from Teacher- and Student-Centered Teaching Methods," *Journal of Higher Education*, 29:456-60, 1968.

Williams, H. Y., Jr. "College Students' Perception of the Personal Traits and Instructional Procedures of Good and Poor Teachers," *Dissertation Abstracts*, 27 (11):3644-A, 1967.

Woodburne, L. S. "Suggested Criteria for Academic Promotion," *AAUP Bulletin*, 29:2, April, 1943.

———. "The Qualifications of Superior Faculty Members," *Journal of Higher Education*, 23:377-82, 1952.

Yamamoto, K. (ed.) *Teaching: Essays and Readings*. Boston: Houghton-Mifflin, 1969.